Near Infrared Spectroscopic Monitoring and Control of a Whole-Cell Biocatalytic Process.

Paul A. Bird, BSc. MRes. AMIChemE

A thesis submitted for the degree of
Doctor of Philosophy to the
University of London

Department of Biochemical Engineering
University College London
Torrington Place
London, WC1E 7JE

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

Accurate and robust monitoring of product and reactants in a complex bioconversion stream is essential for the development of effective process control strategies. To monitor a microbially-catalysed Baeyer-Villiger bioconversion of a cyclic ketone to an optically pure lactone, a near infrared (NIR) spectroscopic method has been developed. The reaction, catalysed by cyclohexanone monooxygenase from *Acinetobacter calcoaceticus* (expressed in *Escherichia coli*) is characterised by substrate (ketone) and product (lactone) inhibition at relatively low concentrations. Quantitative multivariate calibration of a NIR spectrophotometer for ketone and lactone resulted in a standard error of prediction (SEP) at-line of 0.088 and 0.110 $g.l^{-1}$ and on-line of 0.130 and 0.180 $g.l^{-1}$, respectively. The directed modification of quantitative models, by the inclusion of spiked process samples improved the SEP for lactone prediction where bioprocess development meant existing NIR models were not relevant. The monitoring and control of the Baeyer-Villiger bioconversion by NIR has allowed intermittent feeding of ketone such that the concentration of substrate does not rise above 0.6 $g.l^{-1}$. Using this feeding strategy 5.7 $g.l^{-1}$ of lactone product has been produced. This represents a 2-fold increase in productivity. The application of a technique to monitor analytes at low concentration demonstrates the utility of NIR for control of biotransformation processes.

Acknowledgements

Firstly, I would like to thank my supervisor Prof. J.M. Woodley for all his advice, support and encouragement during my time at UCL. In addition, I gratefully acknowledge the grants provided by Pfizer Global Research, the Centre for Scientific Enterprise, London (CSEL) and the BBSRC.

I would also like to thank my industrial supervisor David C.A.Sharp (Pfizer Global Research) for the contribution he made on my behalf. During my time at UCL's Advanced Centre for Biochemical Engineering I had the privilege of working with a number of people, including Dr.Steve Doig, Dr. Chis Baldwin, Dr.Jiang Qui, Dr. Bing H Chen, and Helen E M Law. I am grateful to them and all the people within the Department of Biochemical Engineering who have helped me during my research.

Finally, I would like to thank my family for all their support, especially my Wife for her endless patience and encouragement.

Dedicated in loving memory of my father,
John R Bird (1938-1996).

Comment

"We are not here in this world to find elegant solutions, pregnant with initiative, or to serve the ways and modes of profitable progress. No, we are here to provide for all those who are weaker and hungrier, more battered and crippled than ourselves. That is our only good and great purpose on earth, and if you ask me about those insoluble economic problems that may arise if the top is deprived of their initiative, I would answer, to hell with them. The top is greedy and mean and will always find a way to take care of themselves. They always do."

Words of Michael Foot during the 1983 General Election.

Contents

List of Figures

List of Tables

Nomenclature

A	Absorbance ($\log A_{max}$)
c	Concentration (µmol)
C_{aq}	Equilibrium solute concentration in water (gl^{-1})
CER	Carbon evolution rate ($mmoles.l^{-1}.h^{-1}$)
C_e	Equilibrium solute concentration in ethyl acetate (gl^{-1})
di	Impeller diameter (m)
E	Average energy dissipation per unit mass (W/Kg)
E_r	Enantiomeric ration (-)
K	Ethyl acetate/ water partition coefficient (-)
l	Light path length (cm)
N	Speed of agitator (s^{-1})
OTR	Oxygen transfer rate ($mmoles.l^{-1}.h^{-1}$)
OUR	Oxygen uptake rate ($mmoles.l^{-1}.h^{-1}$)
Pg	Impeller gassed power requirement (watts)
P*ug*	Impeller ungassed power requirement (watts)
Po	Power number (-)
Re	Reynolds number (-)
RQ	Respiratory Quotient (-)
S	Substrate concentration (mmol)
SEC	Standard error of calibration (gl^{-1})
SEP	Standard error of prediction (gl^{-1})
t	Time (minutes)

Ts	Tip speed (s^{-1})
U	Units of enzyme activity ($\mu mol.min^{-1}$)
V	Working volume of the STR (m^3)
ΔP	Density ($Kg.m^{-3}$)
ε	Extinction coefficient ($\mu molml^{-1}cm^{-1}$)
v	Rate of reaction ($mmol.L^{-1}.h^{-1}$)
ν	Kinematic viscosity, ($m^2.s^{-1}$)
γ	Shear rate (s^{-1})

Abbreviations

ACS	Acetyl-CoA synthetase
ADP	Adenosine-5'-diphosphate
API	Active Pharmaceutical Ingredients
ATP	Adenosine-5'-triphosphate
CHMO	Cyclohexanone monooxygenase
CoA	Coenzyme A
CS	Citrate synthase
DOT	Dissolved oxygen tension
DSP	Down stream processing
ee	Enantiomeric excess
FID	Flame ionisation detector
GC	Gas chromatography
GK	Glycerokinase
HPLC	High performance liquid chromatography
IR	Infrared
ISPR	*In situ* product removal technique
NIR	Near infrared
L-LDH	L-lactate dehydrogenase
L-MDH	L-malate dehydrogenase
MLR	Multiple linear least squares regression
NAD	Oxidised nicotinamide-adenine dinucleotide
NADH	Reduced nicotinamide adenine dinucleotide

NADP	Oxidised nicotinamide adenine dinucleotide phosphate
NADPH	Reduced nicotinamide adenine dinucleotide phosphate
OD	Optical density
PCR	Principal component regression
PEP	Phosphoenolpyruvate
PK	Pyruvate kinase
PLC	Programmable logic controller
PLS	Partial least square analysis
PRESS	Prediction Residual Error Sum of Squares
R	Reflectance
RO	Reverse osmosis
SD	Standard Deviation
SE	Standard Error
S/N	Signal to noise ratio
STR	stirred tank reactor
T	Transmission

1. Introduction

1.1. Biotransformations

1.1.1 Enantiomers and Diastereomers

A biotransformation reaction uses a biocatalyst, either as an isolated enzyme or a whole cell, to catalyse a specific chemical reaction in which defined substrate(s) are converted to defined product(s) (Lilly, 1994).

Academic study of biotransformations has developed out of the needs of industry for enantiomerically pure materials. Biotransformations enable the conversion of achiral compounds into enantiomerically pure chiral compounds. Conversion resolution techniques include: asymmetric synthesis, conventional resolution techniques and utilisation of the chiral pool (Kelly, 1998).

A molecule is chiral if it cannot be superimposed on its mirror image. Chiral molecules, because of their handedness, have two forms that differ only in their spatial arrangement. Lactic acid is an example of a molecule (Figure 1.1), which exhibits chirality, the two forms are known as enantiomers. Stereoisomers that are not geometric isomers are called optical isomers. Enantiomers constitute one way in which optical isomers can occur (Figure 1.2).

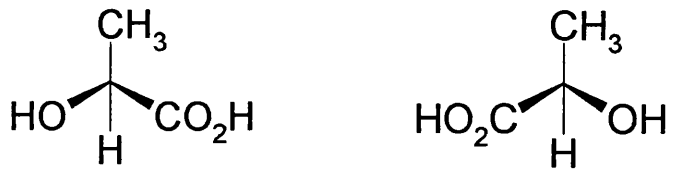


Figure 1.1: Lactic acid exhibits molecular chirality.

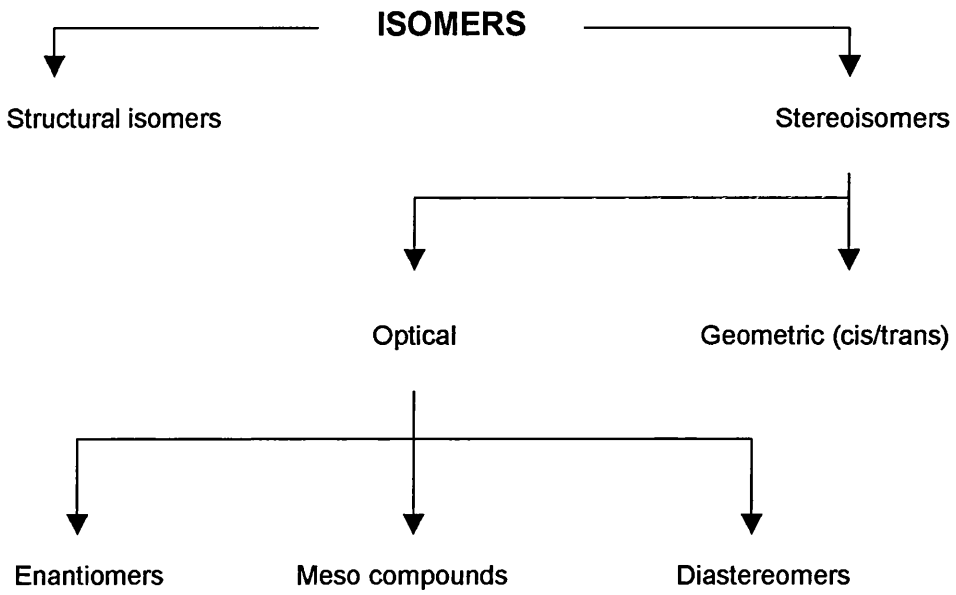


Figure 1.2: The relationship of various kinds of isomers

With the exception of chirality, the structure of a pair enantiomers is the same. This means almost all of their physical and chemical properties are the same. For a single pair of enantiomers there are two sets of properties, which differ:

- Interactions with other chiral substances, and
- The direction of rotation of the plane of polarisation of plane-polarised light.

The difference in optical properties of enantiomers of a molecule is used to distinguish the two forms.

To describe the absolute configurations of a pair of enantiomers the Cohn-Ingold-Prelog system is preferred (Fessenden *et al.*, 1982). This system, also known as the (R) and (S) system describes actual arrangements of groups around a chiral carbon. The letters D and L denote configuration based on

Fischer projection formulas. The R and S system has superseded the Fischer projection nomenclature. The exception being the description of carbohydrates and amino acids.

A mixture containing equal amounts of the two enantiomeric forms of a molecule is known as a racemic mixture. A racemic mixture has no optical activity. Under some circumstances different proportions of the enantiomers will be present in a mixture. The proportion of the enantiomers in the mixture is described by the enantiomeric excess (ee)(Equation 1.1).

$$ee = \frac{A - B}{A + B}$$

Where,
A = proportion of most abundant enantiomer
B = proportion of least abundant enantiomer

Equation1.1: Definition of Enantiomeric Excess (ee):

In, addition to enantiomers, a molecule can have diastereomic forms. Diastereomers arise when there is more than one chiral centre in a molecule, or where the molecule is non-linear. Unlike enantiomers, diastereomers have different physical properties. An example of two pairs of enantiomers, each pair being a distinct diastereomer of the other is 2,3,4-trihydroxybutanal (Figure 1.3).

There are also regioisomers (structural isomers); molecules which have the same molecular formula (isomers) but different regional order of attachment of their atoms. Two regioisomers are therefore distinct molecules (Figure 1.4).

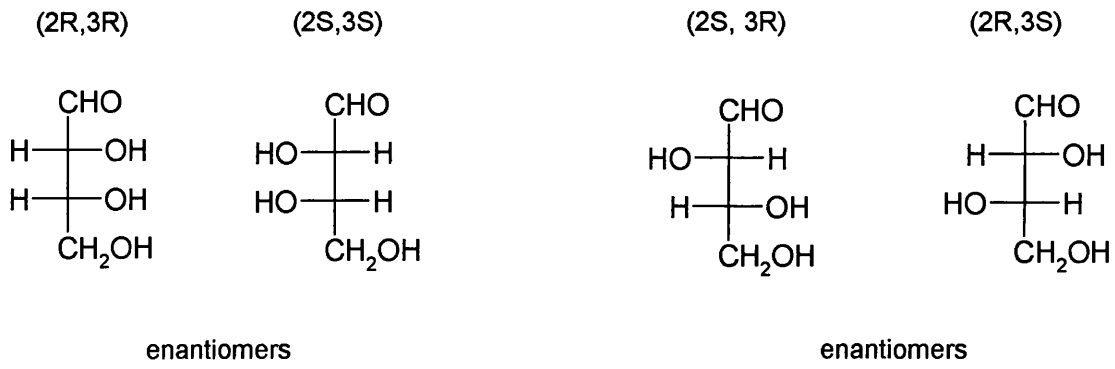


Figure 1.3: The four stereoisomers of 2,3,4-trihydoxybutanal. A pair of stereoisomers which are not enantiomers are called diastereomers.

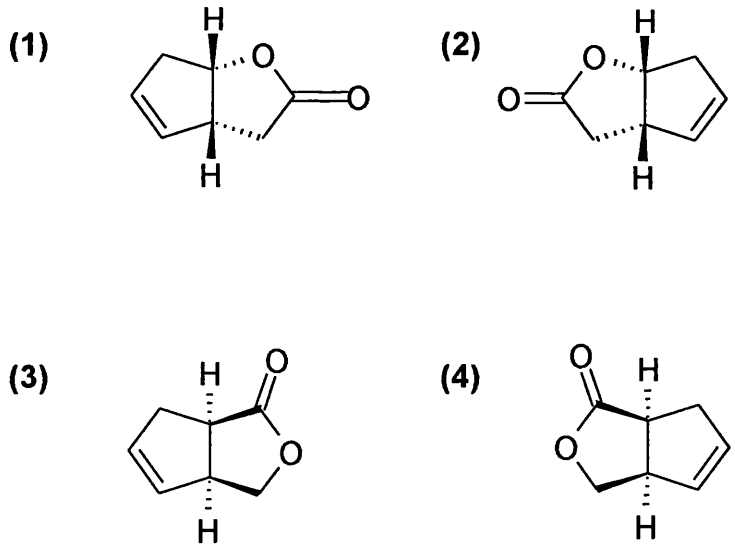


Figure 1.4: The two pairs of lactone enantiomers (1,2) and (3,4) have the same molecular formula but are not stereoisomers (i.e. diastereomers). The different regional configuration of the functional groups means 1 and 3 for example are regioisomers.

1.1.2 Production of Optically Pure Material

There are a number of strategies available for obtaining single-isomers (Figure 1.5). For instance at the point in a process of introducing chirality into a molecule the options include:

- Utilisation of the chiral pool, where the configuration of the molecule originates from the asymmetric configuration of the starting material.
- Resolution of racemates. The required isomer is separated at the end of the reaction. In some cases, the unwanted isomer can be recycled back into the racemate starting material or have its configuration converted to the required isomer.
- Asymmetric synthesis involves introducing asymmetry into a non-chiral material so that a single isomer is formed.

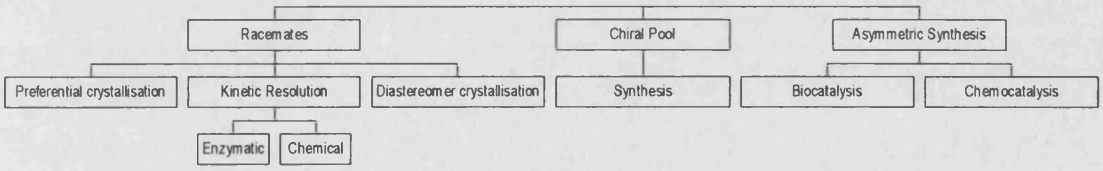


Figure 1.5: Routes to pure stereoisomers. Adapted from (Sheldon, 1996).

1.1.3 The Chiral Pool

The chiral pool refers to optically active compounds, which are available from natural sources. These compounds are either isolated directly from natural sources or manufactured through bioprocesses. They are typically inexpensive and are often produced in significant quantities, 10^2-10^5 tonnes pa (Federsel, 1993).

Chiral pool compounds are used as building blocks, incorporating into the target structure the desired chiral features. Examples of commercially

available chiral building blocks include amino acids, hydroxy acids, carbohydrates, terpenes and alkaloids.

1.1.3.1 Amino Acids

Amino acids are the oldest source of optically active compounds with crystalline glutamic acid isolated from gluten hydrolysate in 1866. Monosodium L-glutamate is manufactured on a large scale with a capacity of 350,000 tonnes pa (Harper, 1983). Both L- and D-amino acids are in demand, for instance D-alanine is required in the sweetener alitome whilst demand for the intermediate L-phenylalanine has also grown.

1.1.3.2 Hydroxy Acids

Hydroxy acids such as lactic acid and malic acid are both available by fermentation. Malic acid is also available by asymmetric synthesis (Wynberg *et al.*, 1982).

1.1.3.3 Carbohydrates

The long chains of carbohydrates (containing many chiral carbons) have meant that they have found limited commercial use as chiral building blocks. Use of them necessitates costly transformations of the long carbon chains to smaller, more useful species. There are examples of industrial carbohydrate chiral building blocks, including D-glucose in the ascorbic-acid producing Reichstein-Grussner process. This process produces >35,000 tonnes pa of L-ascorbic acid and D-sorbitol (Crosby, 1991).

1.1.3.4 Terpenes

Terpenes are characteristic of many varieties of plant life (odorants) and are also important to the pharmaceutical industry. Some of the available, optically active terpenes are: (+)-limonene, (-)-menthol, (-)-carvone, (+)-camphor, (+)-3-carene, (-)-β-pinene, (+)-α-pinene, (-)-α-phellandrene. Terpenes are used mainly as resolving agents or as a source of chirality for asymmetric synthesis, rather than as chiral building blocks.

1.1.3.5 Alkaloids

Alkaloids, which are optically active and mostly dextrorotatory, are organic bases and include a number of highly active materials, e.g. morphine, caffeine, nicotine, atropine, cocaine. Many alkaloids are used in the resolution of racemic compounds on a small scale by crystallisation. Large-scale applications require, due to economic factors, that the alkaloid is recycled from the resulting salt species.

1.1.4 Resolution of Racemates

1.1.4.1 Classical Resolution

Classical resolution was used by Louis Pasteur during the separation of racemic sodium ammonium tartate and is still widely used industrially today. The method uses crystallisation, but not of the enantiomers which have the same chemical and physical properties and so rarely crystallise separately. Instead, a pair of enantiomers undergo a reaction with a chiral reagent or chiral catalyst to obtain a pair of diastereomers. As diastereomers have different physical properties, they can be easily separated by crystallisation. The methodology, principles of the technique, and criteria for good resolving agents have been described by Jacques and co-workers (Jacques *et al.*, 1981), whilst examples of classical resolutions and factors involved in the commercial realisation of such a process have been discussed by Crosby (Crosby, 1991).

1.1.4.2 Kinetic Resolution

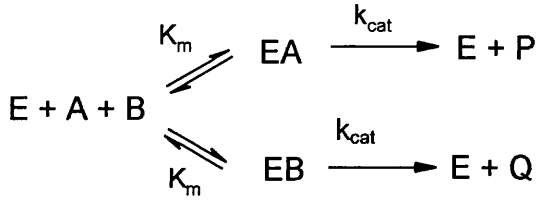
Kinetic resolutions are stereospecific as one of the enantiomers of a racemate is more readily converted to product than the other. In order to achieve this preferential conversion of one enantiomer, chemical catalysis or enzymatic methods are employed. The kinetics of a stereospecific reaction for conversion of a racemate (A,B), which is catalysed by an enzyme, are described by equation 1.2.

The individual rates (Equation 1.3) of the two reactions will depend on the size of the two potential energy barriers. The kinetics are dependent on the substrate concentrations which changes as the reactions progress.

To describe the efficiency of resolution the enantiomeric ratio (E_r) is used. The term E_r is simply the ratio of the two reaction rates (equation 1.4). Provided the reaction is irreversible, the resultant enantiomeric excess (ee) in the product will remain unchanged.

Hence, if the E_r of a stereospecific resolution is high (>300) then it can be said that only the one enantiomer would be transformed and so the reaction would stop at 50% conversion (Sheldon, 1996). Crosby (1991) states that an Er of >20 is sought for commercial processes such that a high ee (98%) can be obtained at <60% conversion. When E_r is very high (>50) the product, if chiral, is simultaneously obtained with a high ee. However, product ee deteriorates as conversion increases. The specificity of an enzyme results in a high E_r value compared to chemical catalysts.

There are numerous examples of industrial enzymic kinetic resolution processes, many of which are discussed by (Crosby, 1991, Liese *et al.*, 1999b, Pandey *et al.*, 1999). The best examples are lipase-catalysed resolutions (Anderson *et al.*, 1998, Jaeger *et al.*, 1998), which are an established synthesis route for the production of chiral carboxylic acids (Tosa *et al.*, 1995) and enantiomerically pure (R) and (S) amines (Balkenhohl *et al.*, 1997, Ladner *et al.*, 1999). These kinetic resolutions are limited by enzyme stereospecificity to a maximal yield of 50%. To overcome this drawback a number of solutions have been developed, including *in situ* racemisation of the unwanted enantiomer (Ebbers *et al.*, 1997, Larsson *et al.*, 1997, Stecher *et al.*, 1997, Um *et al.*, 1998).



Where:

E = Enzyme

AB = A pair of enantiomers A and B

P & Q = Products

K_m = Michaelis constant

K_{cat} = Catalytic rate constant

Equation 1.2: The kinetics of a stereospecific reaction for a racemate which is catalysed by an enzyme.

$$v = \frac{k_{cat}}{k_m}[E][S]$$

V = Rate of reaction ($mmol.L^{-1}.h^{-1}$)

S = Substrate concentration (mmol)

E = Enzyme concentration (mmol)

Equation 1.3: Enzyme reaction rates.

$$E_r = \frac{V_1}{V_2}$$

E_r = Enantiomeric ratio (-)

Equation 1.4: Enantiomeric ratio.

1.1.5 Asymmetric Synthesis

Asymmetric synthesis is the production of enantiomerically pure compounds through transformation of prochiral[1] substrates. Mechanistically the transformation from a prochiral compound is biased in preference to one stereoisomer product. This reaction always involves an optically active agent acting either stoichiometrically or catalytically (including enzymes). This has the advantage that potentially all the material can be realised as the required isomer.

The lyase class of enzymes is used during asymmetric synthesis of amino acids from the corresponding α-keto acid precursors. Aspartase for example is used during the manufacture of L-aspartic acid (Schulze *et al.*, 1999). The use of *Saccharomyces cerevisiae* as a whole cell to carry out asymmetric ketone reductions is a popular route to optically pure alcohols (Griffin *et al.*, 2001, Stewart, 2000).

1.1.5.1 Non-Enzymatic Methods

There are many examples of non-enzymatic methods including hydrogenation, reduction, oxidation, isomerisation, etc. (Crosby, 1991). Crosby has discussed a few examples and makes the point that chemical catalysts lack diversity with respect to the possible substrates accepted.

In addition, stereoselective synthetic transformations require an auxiliary e.g. a catalyst or ligand that introduces or results in asymmetry in the target molecule. The ability of a compound to function as a chiral auxiliary is dependent on an number of requirements (Federsel, 1993):

---

[1] Prochiral: molecules where the addition or substitution of a functional group results in a chiral centre.

- The auxiliary must be available in an enantiomerically pure form (>99% ee) to guarantee a high stereoselectivity is obtained.

- As an auxiliary is used in stoichiometric amounts, it should be quite inexpensive and readily available as well as being recoverable.

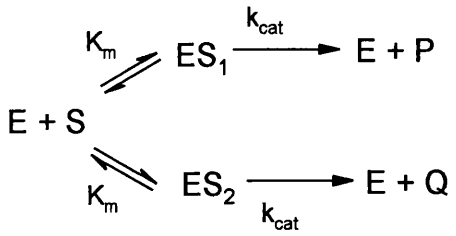
- The auxiliary should be chemically inert under the conditions of the reaction and easily cleaved once the transformation has been completed.

Synthetic procedures are often able to apply favourable reaction conditions to effect chiral induction as the catalyst or auxiliary reagent is more likely to be stable than an enzyme or whole-cell. However, a critical drawback of the chemical asymmetric synthetic route is the low levels of optical purity (Federsel, 1993) that are often observed. The reactions often use a transition metal in conjunction with the chiral auxiliary as an inducer. Many transition metals are toxic and so their presence in the product would have to be reduced to ppm levels in any subsequent processing.

1.1.5.2 Enzymatic Methods

Processes producing enantiomerically pure compounds (McCoy, 1999) are increasingly utilise enzyme-catalysed steps. An enzyme can be used in an isolated state or as a whole cell (micro-organism).

There is a range of enzyme classes: oxidoreductases, transferases, hydrolases, lyases, isomerases and ligases. These enzyme catalysed reactions are usually very selective i.e. where there is only one achiral substrate for the enzyme then the reaction will be selective for one of the pair of enantiomers (enantioselective).An enantioselective reaction can also be described by the rate constants K_{cat}/K_m and the relative free enzyme/ substrate concentrations (Equation1.5).



Equation 1.5: Enantioselective reaction kinetics.

1.1.5.3 Biotransformations Compared to Chemo-Catalysis

A large number of enzymes, which can be applied to single or multi-step biotransformations, exist. There are many more which have not yet been discovered or commercialised. In addition, the ease of isolation of genes coding for target enzymes and the genetic manipulation, evolution and recombination of them has increased still further access to the diversity of enzymes, which can be produced in the amounts required for large scale processes (Bornscheuer *et al.*, 2001, Lilly, 1994, Schmidt-Dannert, 2001, Stemmer, 1994, Zhao *et al.*, 1998).

It is the diversity of enzymes that provides the greatest advantage compared to chemical catalysts. This is because enzymes can often introduce a functional group (and chiral centre) at otherwise unreactive positions. Sometimes conventional (non-enzyme) methods can produce the stereoisomers but often not without the use of environmentally hazardous chemicals (Bull *et al.*, 1999). Biotransformation steps may be used to catalyse reactions that are difficult or impossible to perform using chemical catalysts, or they may demonstrate significant process benefits in a manufacturing operation (Cheetham, 1994). Biotransformations can also be used to carry out conversions that would otherwise require difficult or multiple synthetic chemical protection and de-protection steps. These are often difficult and reduce the atom efficiency of a process (Sheldon, 2000).

Enzymes used need to be produced by way of a fermentation process. This can be expensive, especially if the enzyme is to be used in an isolated form. Production of an isolated enzyme involves purification of the enzyme from the fermentation broth. This typically requires, homogenisation, separation, purification and concentration unit operations. An example is the purification of Penicillin G acylases (PGA). Erarslan and co-workers (Erarslan *et al.*, 1991) purified the enzyme 350-fold by disrupting the cells, removing the cell debris by centrifugation and the nucleic acid material by precipitation. The supernatant fluid from precipitation was further precipitated, the precipitate was subsequently collected by centrifugation. Finally the precipitate was dissolved in phosphate buffer, dialyzed and applied to a DEAE-cellulose column. Specific enzyme fractions from the column were collected. Such a biocatalyst manufacturing process represents a significant cost to the overall biotransformation (Parmar *et al.*, 2000). Further steps are therefore taken to optimise the stability of the enzyme including the use of immobilisation technology (Shewale *et al.*, 1989). Other disadvantages of enzymes are their limited stability, the propensity for inhibition (and/ or toxicity for whole cells) by substrate/ product (Held *et al.*, 1999) and the requirement for regeneration of the enzyme (and or co-factors for isolated enzymes) (Alphand *et al.*, 2003). Enzyme specificity, an advantage of enzyme-catalysed reactions, can under certain circumstances act as a disadvantage. For instance, the adaptation of a developed process for new substrates could be problematic without substantial changes to the enzyme.

The advantages and disadvantages of biotransformations compared to chemo-catalysis are given in Table 1.0. Each application of chiral technology will have its own specific requirements and priorities. The decision to use biocatalysts or not has been discussed by Tramper (1996) who concluded that at present, experience indicates that only when most or all odds are greatly in favour of use of a biotransformation, or when there is no chemical alternative, will it be the process of choice. Thus to transform prochiral substrates or favourably resolve a racemate using an enzyme requires all aspects of a biotransformation to be addressed.

Advantages	Disadvantages
• Reduced number of reaction steps • Fewer and (sometimes) cheaper reagents • High stereoselectivity, regioselectivity, and chemoselectivity. • Milder reaction conditions • Number and diversity of enzymes available • Specific in action • Can introduce functional group at otherwise unreactive positions • Can be applied as a green technology	• Lack of process knowledge and experience • Poor integration of biotransformations step into overall process • Biocatalysts can be expensive to produce. • Enzymes are often unstable and so are not re-usable • Absolute substrate specificity may make the process inflexible. • Low volumetric productivities are common.

Table 1.0 Advantages and Disadvantages of biotransformations compared to chemocatalysis (adapted from Cheetham, 1994).

1.1.6 Significance of Chirality to Industry

Industrial biotransformations cover a wide range of applications, from the use of enzymes as constituents of detergents and analytical tests, the application of biotransformations in the manufacture of active pharmaceutical ingredients (API's), to their use in the paper and textile industries and their application in the brewing and dairy industries (Tramper, 1994). A number of papers review the application of biotransformations to industrial manufacture (Bommarius *et*

al., 1998, Bruggink *et al.*, 1998, Fessner, 1999, Liese *et al.*, 1999a, McCoy, 1999, Ogawa *et al.*, 1999, Schmid *et al.*, 2001, Schulze *et al.*, 1999, Shanley, 1998, Sheldon, 1990, Stinson, 2000, Yagasaki *et al.*, 1998). Over 150 processes to date have been implemented in industry (Liese *et al.*, 2000), and increasing number of which are environmentally clean (Bull *et al.*, 1999, Sheldon, 1996, Sheldon, 1997).

There has been an expansion in the number of pharmaceuticals produced as single stereoisomers over the last decade, a trend which looks set to continue (Borman, 1990). In 1999 single stereoisomer drugs represented 32% of the worldwide pharmaceutical market, up 16% from 1998, with sales of $115 billion (Stinson, 2000). Enantiomers can differ considerably in their pharmacological effects. However many licensed drugs are racemates, and evidence is growing that a mixture of two enantiomers can cause complications when each enantiomer exhibits differences in biological activity (Richards *et al.*, 1997, Stinson, 2000). The reason being that the body's receptors for a given drug are often chiral. This results in the formation of diastereomeric complexes between chiral drugs and drug receptors. It is the difference in physical properties of the two-diastereomeric complexes that causes the observed differences in biological activity of a pair of enantiomers.

Drugs that are racemates can be classified into groups depending on what kind of activity and potency they exhibit:

- Both enantiomers have identical (or nearly identical) pharmacological activity and potency.
- The enantiomers have similar pharmacological activities but different potencies.
- The isomers have qualitatively different pharmacological activities.
- All activity resides in one enantiomer, and the other enantiomer is essentially an impurity.

Thalidomide serves as an excellent example of a racemic drug. Thalidomide, first synthesized in 1953 as a racemate, was widely prescribed for morning sickness from 1957 to 1962. However, Thalidomide became an athema in 1961 when it was found to be seriously teratogenic (creating malformation in embryos) having caused serious birth defects in more than 10,000 babies (Cheymol, 1965). Thalidomide has a stereogenic carbon atom, and so exists as two enantiomers. Tests with mice in 1961 suggested that only one enantiomer was teratogenic while the other possessed the therapeutic activity. The clinical use of thalidomide was stopped until recently. In 1998, FDA approved the use of thalidomide for the treatment of the debilitating and disfiguring lesions associated with erythema nodosum leprosum, a complication of Hansen's Disease, commonly known as leprosy (Muller, 1997).

Where one enantiomer interacts with a cell receptor to produce the desired outcome and the remaining enantiomer is biological 'ballast' then there are obvious benefits from developing the drug as the single isomer to enhance safety and tolerability. Such an approach may also speed the progress of the drug through regulatory channels.

The latter point has become increasingly important since 1992 when both the US Food and Drug Administration (FDA) and the European Committee for Proprietary Medicinal Products required manufacturers to research and characterise each enantiomer in all drugs proposed to be marketed as a racemate (FDA policy statement for the development of new stereoisomeric drugs, 1992).

Due to some single isomer drugs e.g. the anti-obesity drug Redux, having reduced side effects the FDA has been considering incentives (Richards *et al.*, 1997), for developing single enantiomer drugs, including enantiomers of previously approved racemates.

There is therefore an increasing interest by industry in chiral technology. Pharmaceutical companies have responded to regulatory pressures and growing drug complexity by recognising a distinct capability can be gained

from integrating isomeric purity into process design. Concurrently there is also increasing pressure on drug companies to reduce prices.

These issues provide impetus for industry to re-asses the approach taken to drug production. A drug today will typically need around a dozen steps for its synthesis from basic organic chemicals (Richards *et al.*, 1997). However, the application of enzyme-based catalysts has the potential to reduce the number of steps, and increase the productivity of producing high value complex molecules e.g. Crixivan (Patel, 2001, Schulze *et al.*, 1999), which have multiple chiral centres. This applies particularly to those processes, which requires the use of conventional chemistry requires a number of protection/ deprotection steps to realise the chiral molecule.

Some examples of bioconversions that have become commercial processes include the production of high fructose corn syrup and low lactose milk (Cheetham, 1994), the production of acrylamide (Yamada *et al.*, 1996) and the use of penicillin acylase to produce 6-aminopenicillic acid (Ospina, 1991, Parmar *et al.*, 2000, Schulze *et al.*, 1999). Bioconversions used in the manufacture of API are high value low volume fine chemical intermediates (Liese *et al.*, 1999b). In these cases the complexity of the product often makes a bioconversion route the only realistic option.

The biocatalysts most commonly used are hydrolases, such as lipases (Anderson *et al.*, 1998, Matsumae *et al.*, 1993), amidases (Sonke *et al.*, 1999), and dehydrogenases (Ladner *et al.*, 1999). These biocatalysts are readily available and easily applied at a process scale. Other classes of biotransformations have yet to be fully exploited commercially despite the ability of many of these biocatalysts to perform chemically difficult reactions (Faber *et al.*, 2000). These bioconversions include asymmetric synthesis by decarboxylases (Ward *et al.*, 2000), asymmetric C-C bond formation (Turner, 2000), aldolases (Blayer *et al.*, 1999, Guengerich, 2001) and oxidoreductases (Boyd *et al.*, 2001, Holland *et al.*, 2000). The main reason for the limited application of these bioconversions is because it has been found to be difficult to implement economic processes due to constraints

imposed by the biocatalyst, the reaction substrates and products and interactions between them (Lilly *et al.*, 1996).

1.1.7 Whole-Cell and Isolated Enzyme Biotransformations

Both the chemical and pharmaceutical industries are actively seeking innovation driven solutions for the development of both established and new products. One approach for process chemists is to consider biotransformations as a synthetic tool. Enzymes are being used either in their isolated form or contained within microorganisms to catalyse a considerable number of chemical reactions (Liese *et al.*, 1999b, Liese *et al.*, 1999a, Margolin, 1993, Schulze *et al.*, 1999). Enzymes are typically highly specific and selective, and it is these attributes that can help avoid byproducts such as toxic or carcinogenic intermediates (Li *et al.*, 1998), so simplifying the downstream part of a process.

Where optically pure products are desired then the ability of a biocatalyst to assist in stereo- and or regiochemistry is of very significant interest. Typically, optical purity can be achieved in two ways using biocatalysts: first, by resolving a racemic mixture and secondly via asymmetric synthesis of a prochiral reactant. Both approaches are being pursued in the case of selective oxidation reactions, which are of particular industrial interest. Currently 80% of industrial biotransformations are achieved by a kinetic resolution where the biocatalyst acts stereospecifically on a racemic substrate, either to convert a single enantiomer to the desired product or to remove an unwanted enantiomer from the starting mixture (Schulze *et al.*, 1999).

In such reactions, the biocatalyst is often used as an intact microorganism. The use of whole cells as catalysts has the advantage that cofactors (such as NADPH) can be recycled *in-situ.* For example, D-p-hydroxyphenylglycine is an important intermediate for the synthesis of amoxicillin and other antibiotics. To produce the D-amino acid on an industrial scale the enzymes hydantionase and carbamylase are combined in a recombinant whole cell

E.coli system, which constitutively expresses the two enzymes (Grifantini *et al.*, 1998).

A comparison of whole cell and isolated enzyme biocatalyst is given in table 1.1.

Whole cells	Isolated Enzymes
Cheap to manufacture	Expensive to manufacture
Down stream processing of product complicated due to presence of biomass and side-reaction.	Often require co-factors which can cost more than that of the enzyme itself
	Increased selectivity
Large volume reaction mixtures are required to transform small amounts of material.	Performed in more concentrated solutions
Multiple biotransformation steps	Can be performed in organic solvents,
	Single biotransformation step

Table 1.1: A generalised comparison of whole cell and enzyme biotransformations.

The development of biotransformation processes has exploited technology associated with advances in biotechnology, including:

- Optimisation of enzyme activity by site selection and or mutagenesis
- Heterologous expression of gene of interest to increase productivity of enzyme and or simplify the fermentation process.
- Over expression by incorporation of a suitable promoter into the gene

For instance, the gene for cyclohexanone mono-oxygenase which catalyses Baeyer-Villiger reactions has been transferred from the class 2 pathogen *Acinetobacter calcoaceticus* into a yeast (Stewart, 1997) and independently *Escherichia coli* Top10 [pQR239] (Doig *et al.*, 2003b).

1.2 Near Infrared Spectroscopy

1.2.1 Monitoring Biotransformations

Monitoring a bioprocess on-line using analytical techniques where the data is fed into quantitative or qualitative models could allow a process to be controlled for improvement of productivity (Rani *et al.*, 1999, Riesenberg *et al.*, 1999). Bioprocess variables that are routinely measured on-line include; temperature, impeller speed, aeration rate, pH, pO_2 , composition of outlet gasses (O_2 and CO_2), turbidity, and foam detection.

Microbial cell growth and product formation can be optimised by control of these variables. Erythromycin biosynthesis in batch operation was optimised (Cherug *et al.*, 1979) by evaluating optimal temperature and pH profiles. However, the compositions of bioreactor medium and reactant bioprocess streams are not routinely monitored on-line. This is due to the complexity of on-line sensors, which are often required to provide real-time information, be sterility friendly, of known accuracy and be reliable. This means substrate, product concentrations etc. are typically measured, if at all, off-line using techniques such as HPLC and GC. These assays are often tedious, invasive, requires sample handling and difficult to do in real time. Sample preparartion and analysis typically takes 10s of minutes. Conventional laboratory techniques are therefore not suitable for rapid and effective process control.

Traditional spectroscopic techniques in the UV and visible electromagnetic regions cannot be used easily in a bioprocess environment because the turbid nature of the broths can cause light scattering effects. Off-line analyses such as HPLC and GC are capable of determining the concentrations of by-products and nutrients in cell culture (Christensen *et al.*, 1996, Ozturk *et al.*, 1997). However, they also require samples to be removed from the reactor prior to analysis. Removing samples from a process stream or reactor can cause problems including:

- Bioreactor contamination
- Limited measurement frequency

- Complicates automation schemes
- Poor response times

Despite these problems, significant progress has been made in the design of sampling systems and devices (Christensen *et al.*, 1996).

The lack of any on-line information concerning substrates, products and the biocatalyst during a biotransformation delays a fast response, which is required for control. The same is true for metabolites in a fermentation broth. Usually off-line data have large sampling times, and thus the number of process data points is limited. This dictates the number of meaningful parameters that can be estimated. A rapid and more direct method of analysis would yield more timely process information and improved bioprocess control.

Mid-infrared (mid-IR) and near-infrared (NIR) technology has been used on-line or rapid off-line (at-line) to monitor substrate and product concentrations in fermentation processes (Almond *et al.*, 1999, Arnold *et al.*, 2000, Doak *et al.*, 1999, Hall *et al.*, 1996, Pollard *et al.*, 2001, Shaw *et al.*, 1999). The NIR region is ideally suited for the non-destructive analysis of percent levels of components in highly scattering and strongly absorbing matrices, such as fermentation broths (Brimmer *et al.*, 1993). It is the application of this technology to a biotransformation environment which is the subject of this thesis.

Process optimisation requires an advanced level of control. Understanding and monitoring a process can achieve control, allowing for the maintenance of optimum conditions. Suitable on-line sensors are therefore required (Christensen *et al.*, 1996, Galindo *et al.*, 1998, Marose *et al.*, 1999, Schuster, 1999, Vaidyanathan *et al.*, 1999b). Once process information is available, advanced control strategies require the development of mathematical models, which are able to describe the behaviour of processes (Konstantinov *et al.*, 1991, Shioya *et al.*, 1999). Examples of control model programmes include the optimisation of baker's yeast production (Kurtanjek, 1998,

Ramirez *et al.*, 1981) lactic acid production (Shi *et al.*, 1998), and ethanol production (Karim *et al.*, 1997).

1.2.2 Control and Optimisation

When it comes to optimisation of a bioprocess using the existing monitoring and control technologies a gap is often observed between the mathematical model/s and the real process (Karim *et al.*, 1997, Kurtanjek, 1998, Schuster, 1999). The reasons for this discrepancy include the descriptive power of the knowledge inputs i.e. often a complex system is being described by a limited number of factors; the mathematical models are non-linear with time dependent-parameters, and a limited number of operational variables are available in spite of many control objectives. In addition to the lack of appropriate process information the problem of data processing is important. A large amount of data is often collected along with noise. This noise must be minimised by treating the data in some way. This problem is observed also with NIR calibration model building (Blanco *et al.*, 2001, Ding *et al.*, 1999).

Bioprocesses are usually operated in the batch or fed-batch mode, and only rarely in continuous culture mode. Thus, the process state changes dynamically during a microbial cultivation. Similar time-dependent dynamics occur during a whole-cell biotransformation. However as previously mentioned, it is not possible in many cases to directly measure the important state variables such as product concentration or production rate for on-line optimisation. In addition, the various perturbations of initial conditions are rather extensive. Consequently, the sophisticated control systems mentioned are required.

The work reported here aims to develop a chemometrics based quantitative model for one or more bioprocess state variables using NIR as an on-line or at-line sensor. This sensory evaluation of a biotransformation process will yield information that can be exploited by a control system, which, is designed to increase productivity. The nature of the process control is to be determined but may be used for one or more of the following:

- Substrate feed control/ monitoring
- Product rate formation monitoring
- Monitoring control of process state
- Down stream unit operation optimisation

1.2.3 Infrared Monitoring

1.2.3.1 Advantages of NIR and IR Spectroscopy

IR and NIR have considerable scope for bioprocess monitoring and or control. Advantages of the technology include:

- Analysis times under 1 minute are possible on-line (Macaloney, 1996).
- Often, no sample preparation is required and the sample may be re-introduced into the bioreactor. This is advantageous in a process environment where time is an important factor in the analysis (Howard *et al.*, 1998).
- Non-invasive and non-destructive analysis is possible. Flow injection systems and off-line analysis require that the samples be removed from the reactor prior to analysis (Christensen *et al.*, 1996, Ozturk *et al.*, 1997). This removal step can create concerns for bioreactor contamination, and restrict the number of samples taken from a small bioreactor (Riley *et al.*, 1998).
- Cost per analysis is very low (no reagents are used).
- Spectral bands in mid-IR and NIR can have clearly assignable spectral response. For most bioprocesses, spectral bands (especially in the NIR) are overlapped, broad with poor baseline resolution. This is because a typical NIR spectrum encompasses information relating to all the constituents of the sample matrix. Such sensitivity to process variability has been used to investigate bioprocess quality assessment monitoring in a composite manner (Hagman *et al.*, 1998, Vaidyanathan *et al.*, 2001)

allowing for the physiological status of bioprocesses to be monitored at the cellular level (Schuster, 1999).

- Automated correction of background and interference is performed in instruments using computer algorithms. Second derivative pre-processing of spectral data can enhance spectral features and eliminate baseline offset. Aeration and cellular matter in the sample matrix cause light scattering altering the effective sample path-length, producing spectral base-line offset (Shaw *et al.*, 1999). Background information such as that attributable to water (a strong IR absorber) may be subtracted or used as a ratio (Riley *et al.*, 1997, Small *et al.*, 1993), although in some cases this correction has not affected the modelling ability of the algorithm (Yano *et al.*, 1997).

- Detection limits can be very low. The NIR region is ideally suited for non-destructive analysis of fraction of percent levels (<0.1%) of components in highly scattering and strongly absorbing matrices, such as fermentation broths (Brimmer *et al.*, 1993).

- Molecular structural information can be derived from spectra (Pollard *et al.*, 2001, Silverstein *et al.*, 1981).

- Simultaneous monitoring of several chemical species. Monitoring of bioprocesses requires sensors providing real-time measurements of several variables to analyse, model and control optimally the time course of these processes in bioreactor vessels (Acha *et al.*, 2000, Macaloney *et al.*, 1997).

However, the application of the technology to a bioprocess environment and specifically to a biotransformation environment compared to simpler chemical reaction processes is not straightforward. The complexity of the reaction matrix and significant changes in concentrations of substrates, products and by-products with time complicate the design of quantitative calibration models.

1.2.3.2 NIR/ Mid-IR: a comparison

Exposing a molecule to radiation results in it undergoing rotational, vibrational, electronic or ionisation processes. Both mid-IR and NIR measure the absorption or emission of radiation as a consequence of vibrations within a molecule. The electromagnetic waves can be described using wavelength (nm), or wave number ($1/\lambda$ cm^{-1}). The mid-IR spectral region has been defined as 400- 4000 cm^{-1} (2500- 25000 nm). NIR spectra are >4000 cm^{-1} (700 – 2500 nm), and extend to the visible region (Figure 1.6). The upper limit is defined by the extent of all the fundamental molecular vibrations whilst the lower limit is defined by the sampling method, matrix or optics of the instrument. A detailed explanation is given by (Ciurezak *et al.*, 2001, Shaw *et al.*, 1999, Workman *et al.*, 1997).

Table 1.2 compares NIR and mid-IR spectroscopy for application to liquid mixtures.

NIR	mid-IR
780-2500 nm	2500 – 25000 nm
12800 – 4000 cm^{-1}	4000 – 400 cm^{-1}
High energy vibrations	Low energy vibrations
Compatible with optical fibre signal transduction	Only conduit optical fibres can be used
Overtone and combination vibrations	Fundamental vibrations
Low absorbtivities of absorption bands	High signal more resolved peaks
Suitable for analysis of samples containing a high proportion of water	Very high water absorbtivities may result in the signal flooding the detector
Cell path lengths typically 1 – 30 mm	Cell path length typically < 0.5mm
No sample preparation required	Sampling maybe required

Table 1.2: A comparison of NIR and mid-IR spectroscopy applied to aqueous matrixes.

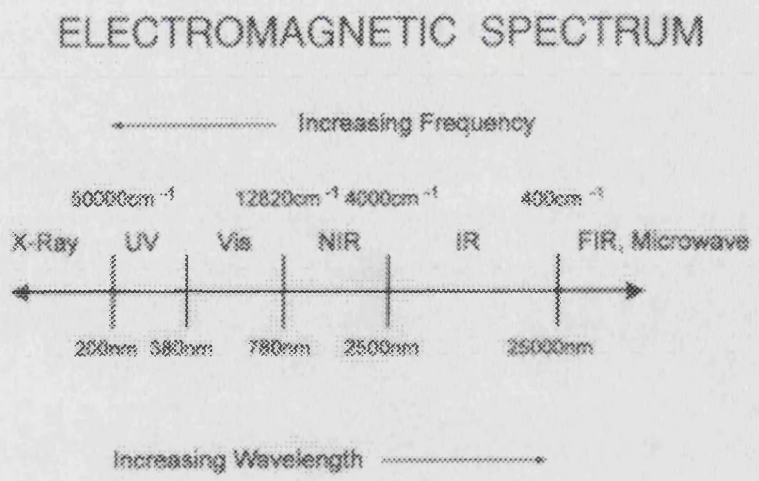


Figure 1.6: The electromagnetic spectrum.

1.2.3.3 Vibrational Energy Oscillations.

Vibrational spectroscopy (mid-IR, NIR) measures energy transitions in a molecule that is exposed to IR radiation. At room temperature, most molecules are at zero energy level. That is, they are vibrating at the least energetic state allowed by quantum mechanics (Ciurezak *et al.*, 2001). These energy transitions are due to molecular vibrations that are analogous to harmonic oscillations. This should not be confused with other spectroscopic methods. UV, visible and x-ray spectra result from excitation of electrons in a molecule or atom respectively. Hooke's law (Equation 1.6) describes an harmonic oscillation. The harmonic oscillation describes well the fundamental vibrational frequency of a diatomic molecule but not polyatomic molecules. In actual molecules, the electron withdrawing or donating properties of neighbouring atoms and groups determines the bond strength and length, and thus the frequency of vibration. It is these species-specific vibrations that determine a substance's characteristic spectra.

$$F = -kx$$

Where,

F= resting force

x= displacement from equilibrium

k= force constant

Equation 1.6: Hooke's law

The energy of an oscillator is discontinuous. Changes in energy content occur by means of transitions between discrete energy levels. Vibrational oscillations caused by absorption by the molecule of quanta of electromagnetic radiation produces transitions in discrete energy states. This means only certain amplitudes of the vibrational motion are possible. For an ideal diatomic molecule, which would behave as a harmonic oscillator, Figure 1.7 depicts potential vibrational energy levels. In reality, real diatomic and polyatomic molecules do not obey Hooke's law and give rise to anharmonicity. Figure 1.7 also describes an anharmonic oscillator where energy levels dissociate. Anharmonic motion occurs for the following reasons:

- As two atoms approach, the repulsion between the nuclei causes the potential energy to rise more quickly than the harmonic approximation.
- When the interatomic distance approaches the dissociation point, potential energy levels off.

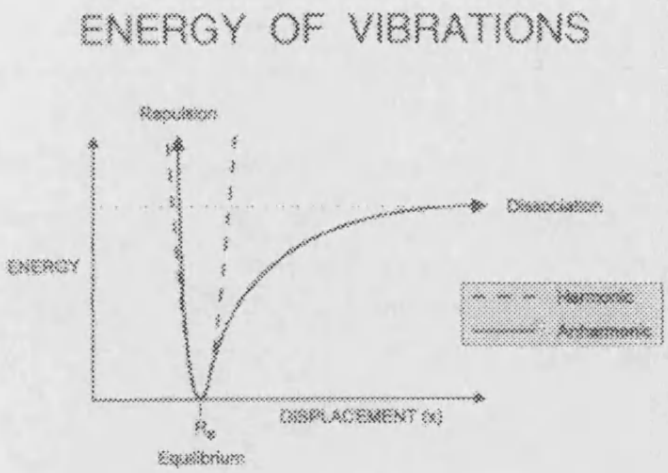


Figure 1.7 : Harmonic and anharmonic oscillations.

1.2.3.4 Fundamental, Overtone and Combinational Molecular Absorptions.

Imparting energy in the form of radiation to a molecule causes molecular vibrations to jump from one energy level to another. Transitions between ground state to energy level one give the fundamental absorption if this leads to a change in the molecular dipole moment. Transitions between ground state to energy level two give the first overtone. Transitions between ground state to energy level three give the second overtone. There are often transitions between multiple states giving rise to combination bands.

Harmonic motion explains fundamental molecular absorption that occur in the mid-IR region. The weaker bands that occur in the NIR region due to anharmonicity are 1-2 orders of magnitude weaker than in the mid-IR and are overtones (Ciurezak *et al.*, 2001). Overtones occur at approximately at ½ and $^{1}/_{3}$ of the fundamental absorption wavelength. However, anharmonicity causes deviations from the predicted.

1.2.3.5 Molecular Vibrations

It is possible to explain molecular vibrations in two ways; stretching (Figure 1.8), a change in inter-atomic distance along the axis of the bond and bending, a change in the bond angle.

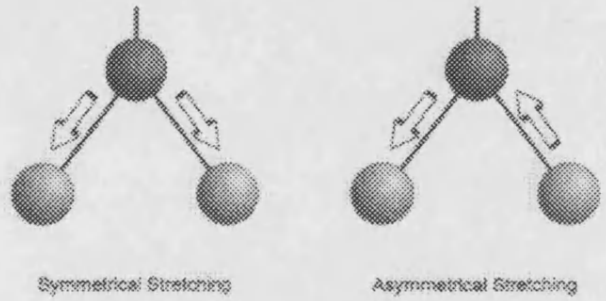


Figure 1.8: Molecular vibrations: stretching; a change in inter-atomic distance along the axis of the bond.

Bond stretching usually has larger vibrational amplitude than bending, so stretching vibrations are usually the strongest. Hydrogen is the lightest atom and will vibrate with the largest amplitude in stretching, this motion deviates most from the harmonic. The intensity of the absorption depends on the degree of anharmonicity.

Many organic functional groups e.g. -OH, -CH, -NH have characteristic frequencies where absorption's are observed. IR spectra can therefore be used to give a characteristic fingerprint to a molecule. The application of mid-IR therefore lends itself to identification of unknowns. Since most of the absorptions are repeated in the NIR as overtones and, combinations of the fundamental vibrations, this region can be used to identify molecules too. However due to spectral overlapping NIR spectra are more often used to

confirm the identity of a species in a matrix which has already been taught to the system.

NIR absorption's are generally 10-100 times weaker in intensity than the fundamental mid-IR absorption bands (Ciurezak *et al.*, 2001). The strong amplitude associated with mid-IR absorption's may require sample preparation due to light scattering or signal flooding of the detector. In contrast, the weakness of NIR absorption bands is an advantage for on-line direct analysis of process samples because the radiation path length can be adjusted in order to detect appropriate signal intensity.

1.2.3.6 Beer-Lambert Law

Vibrational spectroscopy, like other absorption spectroscopic techniques is based on the Beer – Lambert law (Equation 1.7).

$$A = \text{Log } 1/T = \varepsilon c l$$

Where,
A = optical density
T = transmission
ε = molar extinction coefficient ($\mu mol ml^{-1} cm^{-1}$)
l = path length (cm)
c = concentration ($\mu mol ml^{-1}$)

Equation 1.7: Beer – Lambert law

From Equation 1.7 it can be seen that absorbance is proportional to concentration of species and path length of radiation. The Beer-Lambert law only holds in transmission if there is no reflectance or scatter (typical when

analysing solids), the monochromatic light is produced in the same way, and there is no chemical change (Silverstein *et al.*, 1981).

With infrared transmission spectroscopy, the spectrum recorded is a ratio of the radiation intensity emerging from the sample and reaching the detector to the incident on the detector in the absence of the sample

1.2.4 NIR Instrumentation

1.2.4.1 Transmittance or Reflectance

A NIR instrument can be configured for transmittance or reflectance (Figure 1.9). There are many different accessories based on transmittance or reflectance vibrational spectroscopy and designed to deal with the variety of sample matrices. Transmittance is generally applied to opaque liquids where the path-length will not be adversely affected by scattering. Solids (e.g. pharmaceutical tablets) are analysed using reflectance technology (Herkert *et al.*, 2001). The Beer-Lambert law does not hold when the path length ceases to be constant (Silverstein *et al.*, 1981). This may occur due to scattering. Therefore, log 1/T is not useful for most measurements of solids. In such a case log 1/R gives a reasonable approximation to concentration. The term reasonable is employed, as there are problems with path length related to physical properties of solid samples (porosity etc). Bioprocess broths may initially be opaque and suitable for transmission probes. However, as the process progresses an increase in turbidity is often observed. The nature of the matrix will therefore become highly scattering. Under such circumstances a reflectance based NIR probe is used configured to have a relatively large path length e.g. 20mm. This produces an infinite penetration depth i.e. any change in turbidity (penetration thickness) causes no change in the log (1/R) intensity (Vaccari *et al.*, 1994).

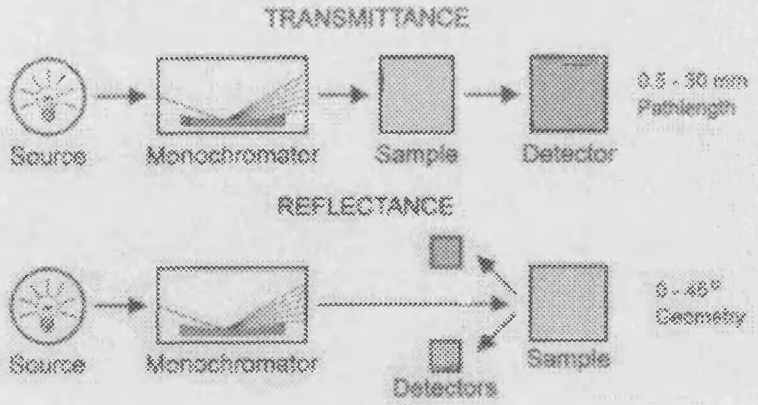


Figure 1.9: Transmission and reflectance configurations of a NIR spectrophotometer.

1.2.4.2 NIR Radiation

The source of monochromatic light in the NIR instrument used during this work was a tungsten-halogen lamp, which is useful in the region of 380 – 3000 nm. The tungsten-halogen lamp is very hot. In contrast, IR sources are incandescent and emit less energy. Lamps found in NIR systems can reach 3000K. The combination of a hot source (NIR radiation) and an appropriate detector allow for high signal to noise ratios (s/n) in the NIR. This is not true for IR instruments, which compensate for lower s/n by the application of Fourier Transform Infrared (FTIR) technology (Hirschfeld, 1983). FTIR probes use the principle of internal reflection to permit analysis of complex, dynamic processes (Doyle *et al.*, 1990, LeBlond *et al.*, 1998, Milosevic *et al.*, 1995, Sheridan *et al.*, 1991).

The absorption of NIR rays in the 700 to 2500 nm region (14,000 cm^{-1} – 4000 cm^{-1}) is related to the overtone and combination bands of the fundamental infrared vibrations of the –CH, -NH and –OH molecular groups. The method is therefore suitable for monitoring any bioconversion where the constituents of interest are organic molecules (Silverstein *et al.*, 1981). Mid-IR absorption

at fundamental frequencies (1500 – 400 cm^{-1}) are typically 10 - 100 times stronger than NIR (Brimmer *et al.*, 1996, Stallard, 1997) and capable of providing distinct spectral features of a compound from a mixture (Pollard *et al.*, 2001). Mid-IR *in-situ* analysis of aqueous mixtures using attenuated total reflection has been demonstrated (Doak *et al.*, 1999, Pollard *et al.*, 2001). The mid-IR system uses optical conduit technology containing knuckles with internal mirrors as a means of guiding the radiation between the sensor and the spectrophotometer.

Since spectral features are broad and overlapped in the NIR region (Ciurezak *et al.*, 2001), resolution of spectral data for analyte calibration therefore requires chemometric techniques for multivariate analysis. In addition, water strongly absorbs IR radiation exhibiting strong OH-bands at 1450 and 1940 nm (Brimmer *et al.*, 1993), making extraction of quantitative information from the NIR spectra more difficult. However, the weak and broad features of absorption bands of NIR when applied to complex aqueous based mixtures are an advantage since no sample preparation or dilution is required. NIR methods use bundles of flexible optical fibres, such that the spectrophotometer can be remotely situated some distance from the reactor. Since the NIR technique is rapid, requires no reagents for analysis and the instrumentation can be located some distance from the reactor site, there is a real potential to improve bioprocess operation in the chemical and pharmaceutical industries by monitoring conversions with NIR spectroscopy.

1.2.4.3 Sample Preparation

It is often necessary to prepare samples before taking vibrational spectra. The choice of sample preparation and NIR or mid-IR accessory is crucial, as the spectrum detail is dependent on this (Hirschfeld, 1983). Sampling techniques as described refer mainly to studies where the chemical and physical nature of solids is of interest. In-order to develop an on-line monitoring NIR technique for a biotransformation process, a method is required that has no sample preparation. NIR spectroscopy does not require sample preparation because the low absorbtivities of bands associated with this region are compatible with moderately concentrated samples and longer

(1-5mm) path-lengths. This allows rapid and non-destructive analysis of samples. It might be the case that on-line samples are sub-optimal for analytical purposes. For instance scattering by cells will affect path length and result in significant scatter around any predicted vs reference regression line (Brimmer *et al.*, 1993). Light scattering effects can be removed by mathematical treatments that can adjust signals for linear and multiplicative effects (Shaw *et al.*, 1999).

1.2.5 Bioprocess Applications of NIR Technology

NIR has only recently been considered as a possible tool in bioprocessing. The reasons for not exploiting the technology have included the cost of the instrument, the lack of mobility of instruments and remoteness of probes, and the complexity of chemometric reduction of data. All these areas have been addressed to produce cost effective, remote systems that are controlled through user-friendly Windows based software (Hirschfeld, 1983).

There are now examples of NIR-based control systems, which allow for fully automated bioprocess production and optimisation. (Gonzalez-Vara *et al.*, 2000) reported a small-scale pilot-plant lactic acid production process, which was fully automated by way of a NIR-based control strategy. This included dedicated software enabling on-line NIR data acquisition and reduction, and automated process management through feed addition, culture removal and/or product recovery by micro-filtration to allow for a continuous fermentation process with recycling of culture medium and cells. The control of feed rates into the reactor was by automatic adjustments of the flow rate of peristaltic pumps. Such developments enable more advanced control systems to be developed.

NIR spectroscopy has been successfully applied in agriculture, food processing, polymeric, chemical and pharmaceutical industries. Numerous NIR spectroscopic methods have been developed including; applications in biological systems (Murayama *et al.*, 1998, Qing *et al.*, 1999), enzyme reaction (Blanco *et al.*, 2000, Chung *et al.*, 1995, Chung *et al.*, 2000), for fermentation processes (Arnold *et al.*, 2002, Arnold *et al.*, 2003), including

fermentations of yeast (Cavinato *et al.*, 1990, Ge *et al.*, 1994), bacteria (Arnold *et al.*, 2001, Vaccari *et al.*, 1994), animal cells (McShane *et al.*, 1998, Yano *et al.*, 1994), and insect cells (Riley *et al.*, 1997, Riley *et al.*, 1998). Wider bioprocess applications include the monitoring and control of a downstream yeast cell homogenate (Yeung *et al.*, 1999). There has been no study of the application of NIR spectroscopy to biotransformations, although real-time monitoring of nitrile bioconversions by mid-infrared spectroscopy (Dadd *et al.*, 2000) has demonstrated the potential of vibrational spectroscopy for the study of biotransformations .

The work to date has mainly concentrated on fermentation steps. There are no published reports concerning NIR monitored whole-cell biotransformations which can have similar physical environments to fermentation reactors.

1.2.6 Chemometrics

NIR absorption bands are broader and can be highly overlapped compared to mid-IR bands. Despite the disadvantage of broad and overlapping bands not distinguishing individual analytes, the advent of chemometrics has meant meaningful information can be extracted from the complex NIR spectra. It has been shown (Brimmer *et al.*, 1993) that a NIR spectroscopic calibration model can be as accurate as the reference analytical technique method used to derive the algorithm.

1.2.6.1 Quantitative Calibration

Prior to a quantitative calibration method the NIR spectra can often be improved by performing a derivative mathematical treatment on the data. First or second derivative treatments can reduce baseline shifts and improve both peak-shape and resolution. However, these treatments can only adjust the data for linear effects (Shaw *et al.*, 1999).

Quantitative spectroscopy involves the construction of empirical relationships between the absorption of NIR radiation and data from a reference analytical technique e.g. HPLC, GC. Any calibration algorithm requires complex chemometrics (Thomas, 1994). Approximating, the intensity of the NIR or

mid-IR spectra is linearly proportional to the number of vibrating species responsible for its occurrence. The standard techniques are:

- Multiple linear least squares regression (MLR)
- Partial least-squares regression (PLS)
- Principle component regression (PCR)

A comparison of the chemometrics techniques has been made by Thomas and Hooland (1990).

MLR is the simplest of the calibration approaches but can only be used when the unique spectral features can be identified for the analyte of interest. Quantitative information can be extracted from single or multiple peak intensities after calibration against standards. However, variations may be present in sampling (e.g. path-length/ scatter) and instrument performance. MLR is less sensitive to subtle matrix variations common to most industrial processes. The NIR region of a multi-component system is likely to contain absorption bands that overlap and in which the compounds have independently variable concentrations. Under such circumstances, MLR is not capable of generating a valid quantitative model.

When unique spectral features cannot be identified as is often the case with multi-component matrices, then PLS and PCR is used. PLS and PCR are full-spectrum methods and can create algorithms for components in complex matrices, where NIR spectral bands are highly overlapped by absorption bands of other components. The crux of producing a robust calibration algorithm for bioprocess broths is to correct for interfering absorptions from other materials and scattering differences that affect spectral baselines and the effective path length of the NIR radiation.

PLS and PCR use a large number of wavelength variables making these methods susceptible to interference's from un-modelled spectral variations compared to MLR. Thus, calibration-training data sets should contain all possible variations. However, in the industrial environment it is difficult to

ensure that all sources of spectral variation have been modelled. Temperature, pressure, viscosity, and other process variables fluctuate during bioprocesses. When vibrational spectra are measured on or at-line for analytical and control purposes, the fluctuations influence the shape of the spectra (Wulfert *et al.*, 2001). MLR models are typically developed from 1 or 2 wavelengths and so are less sensitive to subtle matrix variations.

An overview describing the steps needed to calibrate a NIR spectrophotometer is depicted by Figure 1.10 and described in Osborne *et al.* (2002). Regardless of the calibration method to be invoked, the first step involves collecting a training data-set which should be representative of the population of samples which are expected to be analysed in the future. With regard to NIR analysis, measurements are made without dilution from three to six process operations to provide suitable sample variability (Arnold *et al.*, 2002). The sample data set is split in to two, to give training and validation data sets. Since NIR models describe empirical relationships using reference analytical chemistry, it is important to select a reference method that posses the best analytical performance (Lavine, 2000). The integrity and applicability of the derived calibration model is dependent on the set of data used to create the model. Performance of derived models can be tested using the validation data set before routine analysis takes place. The validation data set should ideally be independent of the training data-set.

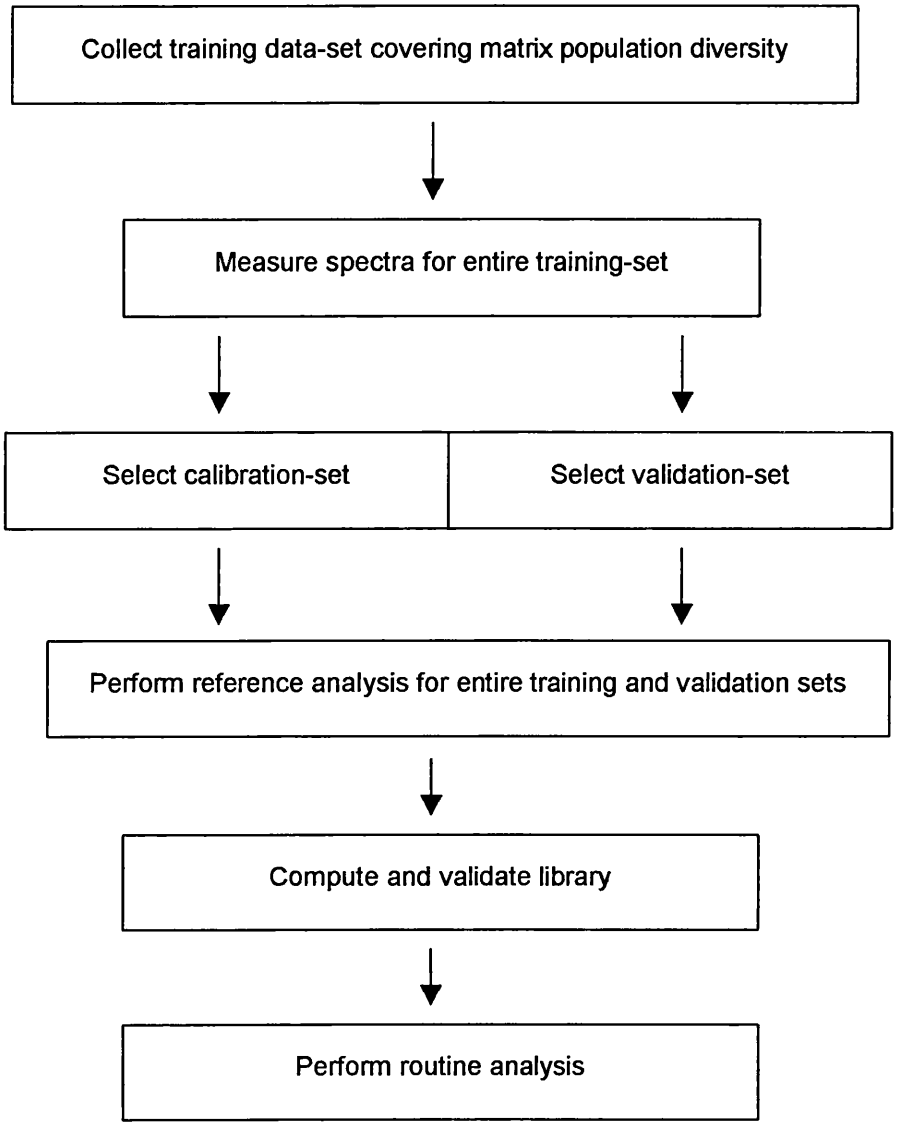


Figure 1.10: Steps required to develop a quantitative or qualitative calibration method.

1.2.6.2 PCR and PLS Calibration Models

PCR and PLS are two commonly applied chemometric techniques used for analysing multivariate vibrational spectroscopic data sets. The use of complex chemometrics has only been possible due to advances in computer hardware and software. However, there are limitations and requirements associated with the technology.

With regard to multi-component systems, the absorbance intensity at a particular wavelength (λ) for an *n* component system may be represented by equation 1.8.

$$\overline{I_\lambda} = K_1c_1 + K_2c_2 + K_3c_3 + ...K_nc_n$$

c_1 to c_n = concentration of component number 1 to *n*

K_1 to K_n = path-length constants of component number 1 to *n*.

Equation 1.8

Equation 1.8 assumes that the contribution to the intensity of the components is additive and therefore linear. Equation 1.9; *x* equals the number of data points over the wave-length range λ_1 to λ_x. The *x* equations can be expressed in matrix form, which can be represented by:

$$\overline{I_{\lambda\, x}} = \overline{I_{\lambda\, xn}}c_n$$

Equation 1.9

With *S* standard samples there will be *S* sets of such matrices, represented as:

$$\overline{I_{\lambda\, s}} = \overline{I_{\lambda\, n}}c_{ns}$$

Equation 1.10

Annotation of spectral data in a matrix form is the basis for PCR and PLS modelling. The detail of each method is not addressed here. The

mathematical procedures are all computer based using chemometrics software for data management and reduction.

1.2.6.3 Qualitative Calibration

The steps taken to develop a qualitative method are the same as for quantitative methods (Figure 1.10). Qualitative techniques are often employed to determine if a sample falls within a defined range of variability. Thus, the quality of material can be gauged. The spectrum of a sample is unique to the composition of its constituents. Other samples of similar composition quality therefore have spectra that are similar to the original spectrum. By comparing spectra, the difference between a defined good and a bad spectrum can be determined. Chemometric qualitative discriminate (sometimes called pattern recognition) models allow unambiguous spectral matching to be performed (Lavine, 2000). The discriminate analysis of spectra can be performed by applying a number of algorithms (Workman *et al.*, 1997).

Characteristic differences between qualitative and quantitative approaches are given in Table 1.3.

Qualitative models	Quantitative models
Determine quality of a sample within a defined range.	Large number of training samples required.
Indicate if a samples are contaminated	Large initial time cost.
Models do not require primary calibration data.	Accuracy limited by accuracy of primary method.
Algorithm's can be determined for very complex mixtures.	Determines the quantity of a specific constituent but not quality.
Simple pass or fail answers	

Table 1.3: A comparison between qualitative and quantitative models

1.3 Aim of the Project

The aim of this project was to provide an evaluation of the potential use of near infrared spectroscopy (NIR) as a means of improving the productivity of biotransformation reactions. The development of a system capable of precise real-time measurements of individual components of a complex bioprocess facilitated the realisation of improved control strategies and therefore improvements to bioprocesses productivity.

The biotransformation used to exemplify NIR during this study was the conversion of a bicyclic ketone to two chiral lactone products by cyclohexanone monooxygenase (CHMO). CHMO is one of a class of potentially important biocatalysts, the Baeyer-Villiger Monooxygenases (BVMOs). The enzyme was expressed heterologously in *E.coli* TOP10(pQR239), a whole-cell biocatalyst. The stereo- and regio- selective biocatalyst exhibits both substrate and product inhibition making the process an ideal model biotransformation.

Monitoring is critical to the success of biotransformations as many continue to be optimised by maximising and controlling the stability of enzymes. NIR technology has the potential to satisfy the requirements for monitoring. It is rapid, reagentless, noninvasive, and suitable for light scattering, inhomogeneous, complex mixtures. NIR is also capable of online monitoring of several components simultaneously.

2. The Model System - Baeyer-Villiger Monooxygenase Biotransformation

2.1. Cyclohexanone Monooxygenase

The bioconversion investigated during this study is the conversion of a bicyclic ketone to two chiral lactone products by cyclohexanone monooxygenase (CHMO). CHMO is one of a class of potentially important biocatalysts, the Baeyer-Villiger Monooxygenases (BVMOs). The Baeyer-Villiger biotransformation was studied as a model system to examine the application of NIR spectroscopy to monitoring and control of biotransformations.

Oxidoreductases compromise the class of enzymes that catalyse oxidation and reduction reactions (Sheldon, 1999). BVMOs are oxidoreductases capable of the nucleophilic oxygenation of ketones and aldehydes in a manner analogous to the established peracid catalysed organic chemical reaction from which they take their name (Baeyer *et al.*, 1899). The application of BVMOs and other oxidoreductases to synthetic organic chemistry is of interest due to the high chemo-, regio-, and stereo- specificity that can be achieved (Fessner, 1999, Patel, 2000).

The monooxygenases employ a range of prosthetic groups including:

- Iron (e.g. cytochrome P_{450})
- Copper (e.g. dopamine β- hydroxylase)
- Pterins
- Flavins and cofactors NADPH or vitamin C.

Flavoproteins are conjugated protein dehydrogenases containing flavin and are involved in biological oxidation reactions. There are three principal flavins found in nature; riboflavin, flavin mononucleotide, flavin adenine dinucleotide

Flavin mononucleotide and flavin adenine dinucleotide (FAD) act as prosthetic groups for flavoenzymes, which have a bright yellow colour and are involved in biochemical redox reactions. In biotransformations, flavoenzymes are used as oxygenases. Oxygenases have been classified into two groups (Kelly, 1998):

- Monooxygenases. These catalyse insertion of one oxygen atom from O_2 into a substrate with the remaining oxygen undergoing reduction to H_2O.
- Dioxygenases. Catalyse the insertion of both oxygen atoms from O_2 into a substrate.

The catalytic cycle of a flavoenzyme has been depicted by Ghishla and co-workers (1978). Kelly *et al.*, (1998) described flavoenzymes as a chiral hydroperoxide as they oxidise among other reactions, ketones to esters and lactones, the so-called Baeyer-Villiger reaction (Figure 2.0). BVMOs are a particular group of monooxygenases which use flavin coenzyme as a cofactor and nicotinamide adenine dinucleotide phosphate (NADPH) as a reductant (Willets, 1997). The general mechanism of Baeyer-Villiger reactions has been discussed by the oxidation of cyclohexanone both abiotically and biologically by (Kelly *et al.*, 1998).

The Baeyer-Villiger oxidation in synthetic organic chemistry is achieved by the use of a peracid as the main reactant. This makes any industrial process risky. Preparation of a peracid implies the use of stoichiometric amounts of a peroxide (essentially 90% hydrogen peroxide) as a primary reactant. Peroxide oxidations such as this are hazardous, running the risk of reactions occurring with explosive violence. If the starting material is chiral the chemical route will produce a racemic mixture because the peracid reagent can attack either face of the molecule (Renz *et al.*, 1999).

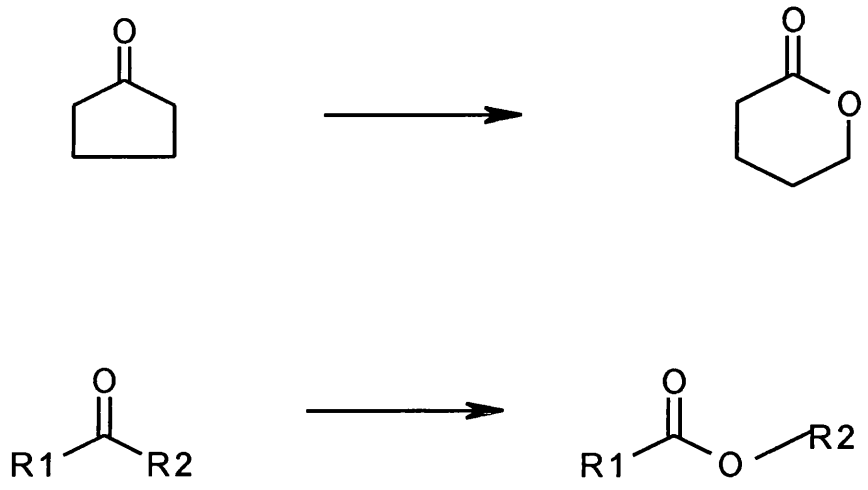


Figure 2.0: The generalised Baeyer-Villiger reaction. The enzyme driven Baeyer-Villiger reaction has an absolute requirement for NADPH and O_2.

A diverse range of microorganisms have been found to produce BVMOs that mimic the chemical reaction (Willets, 1997). Unlike the chemical reaction, BVMO catalysed bioconversions often show high regio- and stereoselectivity. Recently, some limited enantioselectivity has been achieved at low ee values in chemical Baeyer-Villiger reactions by various transition metal based reagents (Corma *et al.*, 2001). The Baeyer-Villiger oxidation of cyclohexanol to adipic acid has been undertaken using: *Nocardia globerula* CL1 (Norris *et al.*, 1971) and *Acinetobacter sp.* NCIMB 9871 (Barclay *et al.*, 2001, Donoghue *et al.*, 1975, Gagnon *et al.*, 1994). The enzyme responsible for this biotransformation was found to be cyclohexanone monooxygenase (CHMO) (Donoghue *et al.*, 1976). CHMO consists of a single polypeptide chain, which binds one FAD as a prosthetic group. The flavoenzyme Baeyer-Villiger activity has an absolute requirement for NADPH and O_2.

Acinetobacter sp. NCIMB 9871 is a class 2 pathogen requiring microbiological containment facilities. However, heterologous expression of the CHMO gene from *Acetobacter sp.* has been achieved in yeast (Stewart, 1997), and *Escherichia coli* Top10 [pQR239] (Doig *et al.*, 2003b, Stewart *et*

al., 2002). The recombinant *E.coli* has been engineered such that the lactone resulting from the Baeyer-Villiger oxidation is not further hydrolysed. Such a reaction (Figure 2.1) is now used as a model chiral biotransformation process.

CHMO catalyses the nucleophilic oxygenation of a wide variety of linear or cyclic ketones to corresponding optically pure esters and ketones (Kelly, 2000, Mihovilovic *et al.*, 2001, Sheng *et al.*, 2001, Zambianchi *et al.*, 2002). CHMO is the best characterised of over thirty BVMOs that have been identified. Optically pure chiral synthons produced via CHMO could be used for the synthesis of a variety of value added fine chemicals and intermediate pharmaceuticals. Recent reviews by Stewart (Stewart, 1998) and Roberts and Wan (Roberts *et al.*, 1998) detail the catalytic range of the CHMO enzyme. A number of models for the CHMO active site have been proposed to explain and predict the specificity of the enzyme (Kelly *et al.*, 1995, Ottolina *et al.*, 1996, Wright *et al.*, 1994, Zambianchi *et al.*, 2001).

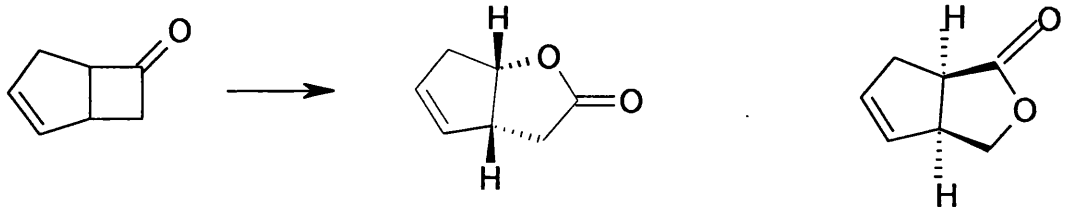


Figure 2.1: The Baeyer-Villiger reaction on bicyclo[3.2.0]hept-2-en-6-one.

BVMOs have been used either as whole cells or isolated enzymes for the transformations of numerous substrates. All BVMOs use molecular oxygen and the reduced cofactor NAD(P)H as a reductant to carry out the Baeyer-Villiger type oxidation (Kelly, 2000, Willetts, 1997). Because these reactions involve NADPH cofactors, the use of whole-cells rather than isolated enzymes is prefered to avoid the need for co-factor regeneration. Isolated BVMOs require an ancillary enzymatic system for the *in* situ regeneration of these expensive cofactors (Walton *et al.*, 2002, Zambianchi *et al.*, 2002).

Such a regeneration system has been carried out (Rissom *et al.*, 1997). Application of BVMOs to the synthesis of natural products is well documented and some key chiral synthons produced via BVMO catalysed bioconversions are shown in Table 2.0. Of specific interest is the production of optically pure lactones, which are chemically highly versatile and are often used as key building blocks for further elaboration of industrially valuable targets and, in particular, biologically active products such as pharmaceuticals, agrochemicals and flavour constituents. However, there are currently no industrial processes, which use a BVMO catalysed step. This is primarily because of low volumetric productivities resulting from bioprocess issues such as product and substrate inhibition and cofactor requirements (Schulze *et al.*, 1999).

Compound	Application	Bioconversion details	Reference
Multifidene	Pesticide	Precursor synthesised from a bicycloketone, using CHMO.	(Carnell *et al.*, 1991)
(R)-(-)- b Baclofen	Neurotransmitter derivative	Precursor synthesised from a substituted cyclobutanone.	(Mazzini *et al.*, 1997)
Sordidin	Pesticide	Precursor synthesised from a 2,6 disubstituted cyclohexanone	(Beauhaire *et al.*, 1995)
Azadirachtin	Insect antifeedant	Precursor synthesised from a bicycloketone, using CHMO.	(Petit *et al.*, 1995)
Sarkomycin A	Antibiotic	Precursor synthesised from a bicycloketone.	(Konigsberger *et al.*, 1994)
Ionomycin	Antibiotic	Precursor synthesised using CHMO.	(Taschner *et al.*, 1991)
Carbocyclic nucleosides	Antiviral agents	Precursor synthesised from norbornanone using CHMO.	(Levitt *et al.*, 1990)

Table 2.0: Chiral synthons produced via BVMO catalysed bioconversions

2.2. A Model Industrial Biotransformation

Biocatalysts are increasingly used during the production of industrially significant compounds (Liese *et al.*, 1999a, Rozzell, 1999, Tramper, 1996). Novel, optically pure products can be synthesised efficiently in a single unit operation by means of a biotransformation due to the characteristically good chemoselectivity, regioselectivity and stereoselectivity of enzymes. Where an alternative chemical conversion is also available for product synthesis, a greater number of reaction steps are typically required. In addition, reaction conditions for biotransformations are often aqueous based, using mild operating temperatures and pressures. Thus, a bioprocess can have convincing advantages with respect to process economics and environmental impact when compared to chemical conversions. However, under operating conditions, the substrates and products are frequently toxic or inhibitory to the biocatalyst. This necessitates the use of advanced control systems (Collins *et al.*, 1995, Collins *et al.*, 1998, Freeman *et al.*, 1993, Hack *et al.*, 2000) requiring process monitoring of the analytes of interest (see section 1.1).

The model reaction (Figure 2.1) is a stereo-selective Baeyer-Villiger oxidation of bicyclo[3.2.0]hept-2-en-6-one to the two corresponding regio-isomeric lactones. The whole cell catalyst, *E.coli* TOP10(pQR239) expressing CHMO from *Acinetobacter calcoaceticus* exhibits regio and enantiospecific oxidation of bicyclo[3.2.0]hept-2-en-6-one (from this point will be referred to as ketone).

The stereoselective oxidation of ketone to its corresponding regio-isomeric lactones (-) 1(S), 5(R)2-oxabicyclo(3.3.0)oct-6-en-3one (ee = 94%) and (-) 1(R), 5(S) 3-oxabicyclo(3.3.0)oct-6-en-2-one (ee = 99%) is achieved by *E.coli* TOP10(pQR239). This reaction has been reported previously (Doig *et al.*, 2002, Doig *et al.*, 2003b). The fed-batch biotransformation was undertaken, in the broth of the preceding fermentation, using an agitated stirred tank (Figure 2.2). From this point lactones (-) 1(S), 5(R)2-oxabicyclo(3.3.0)oct-6-en-3one and (-) 1(R), 5(S) 3-oxabicyclo(3.3.0)oct-6-en-2-one will be referred to as combined lactone.

Under typical bioconversion concentrations, the substrates and products are toxic (Hack *et al.*, 2000) or inhibitory to the bioctalyst (Doig *et al.*, 2003b). The use of a fed-batch stirred tank reactor with feed back control enables the substrate concentration to be maintained in the optimum range. Such a system typically requires monitoring of the substrate at low concentration whilst maintaining a minimum sampling and analysis time. Where a bioconversion is characterised by product inhibition then strategies to circumvent the problem can be adopted (Becker *et al.*, 1997, Collins *et al.*, 1995) or the process can be run to maximise productivity according to the enzyme kinetics (Doig *et al.*, 2003b). The kinetics of the whole-cell Baeyer-Villiger bioconverson of ketone to its corresponding regio-isomeric lactones have been described (Doig *et al.*, 2002). The optimum specific activity of 55 μmol/min/g was observed at a ketone concentration between 0.2 and 0.4 gl^{-1}, whilst at 5 gl^{-1} ketone this was reduced by more than 90%. In addition, when the combined lactone concentration was greater than 4.5 gl^{-1} the specific activity was zero. It is clear that control and optimisation of this bioconversion would be significantly improved by an effective on-line or rapid at-line monitoring technique. Substrate and product inhibition in biotransformations is a frequently observed phenomenon (Table 2.1) making this an excellent representative biotransformation.

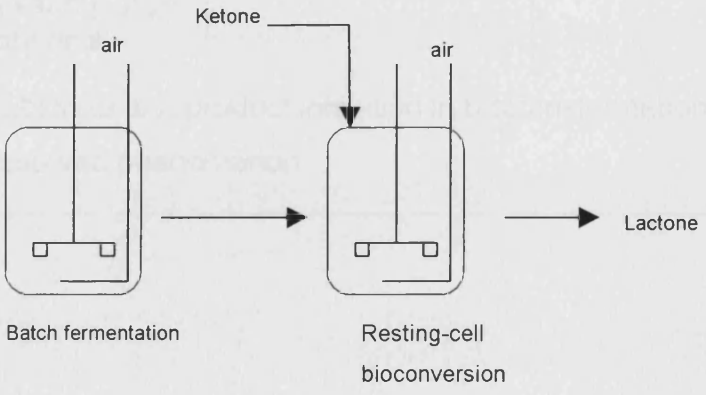


Figure 2.2: The fed-batch Baeyer-Villiger whole-cell biotransformation.

Example reaction	Example biocatalyst	Reference
N-acetyl-D-neuraminic acid (Neu5Ac) condensation	Neu5Ac Adolase	(Blayer *et al.*, 1996, Blayer *et al.*, 1999)
Acrylonitrile conversion	Nitrile hydratase	(Gradley *et al.*, 1994, Graham *et al.*, 2000, Layh *et al.*, 2002)
Synthesis of D-malate	Maleate hydratase	(Michielsen *et al.*, 2000)
Esterification of benzoic acid	Lipase	(Leszczak *et al.*, 1998)
Carbon – carbon bond synthesis	Transketolase	(Brocklebank *et al.*, 1999, Lye *et al.*, 1999, Turner, 2000)
Hydroxylation at the C_3-poaition of a variety of 2-substituted phenols.	Monooxygenase	(Held *et al.*, 1999)

Table 2.1: Substrate and product inhibition in biotransformations is a frequently observed phenomenon.

3. Biocatalyst Production and Biotransformation

3.1. Materials and Methods

3.1.1. Biocatalyst Production

Cyclohexanone monooxygenase (CHMO) was produced during batch growth of recombinant *Escherichia coli* TOP10 [pQR239]. Heterologous expression of the CHMO gene present on pQR239 was under the control of the araBAD promoter induced with L-arabinose. In addition, the plasmid contained a gene conferring ampicillin resistance to the microorganism. *Escherichia coli* TOP10 [pQR239] was a gift from Dr. John Ward, Department of Biochemistry and Molecular Biology, University College London.

Reverse Osmosis water was used in all experimental work. All other reagents were of analytical grade and were obtained from Sigma-Aldrich (Dorset, UK) unless otherwise stated.

3.1.1.1. Storage and Maintenance of the Biocatalyst

Stock cultures of *E.coli* TOP10 [pQR239] were stored at -70°C in 1ml culture aliquots containing 50% sterile glycerol v/v. To prepare stock cultures an individual colony of *E.coli* TOP10 [pQR239] was grown overnight at 37°C on an agar plate containing a complex growth medium with 1.5% w/v agarose and 100 mgl^{-1} ampicillin. This colony was inoculated into a sterile 1.0 l shake flask containing 100 ml of the complex growth medium, to which 100 mg/l of ampicillin had been added via a 0.2 µm sterile filter (Pall, MI, USA). A Stock culture was produced once growth at 37°C in an orbital shaker (Edison House, Hatfield, Herts, UK) rotating at 200rpm resulted in a biomass concentration of approximately 1.0 g $dcw.l^{-1}$. The complex growth medium consisted of 10 $g.l^{-1}$ of glycerol, 10 $g.l^{-1}$ of NaCl, 13 $g.l^{-1}$ of Tryptone Peptone, and 27 $g.l^{-1}$ of yeast extract.

3.1.1.2. Shake Flask Growth

A 1 litre baffled shake flask containing 100 ml of complex growth medium was autoclaved at 121°C for 20 minutes. To this 100 mg/l of ampiciliin was then added via a 0.2 µm filter when the sterile flask had cooled to room temperature. A single 1 ml frozen stock culture was then thawed and transferred into the flask and incubated at 37°C overnight with shaking (250rpm).

3.1.1.3. *E.coli* Top10 [pQR239] Fermentation

Batch fermentation of *E.coli* Top10 [pQR239] to produce quantities of CHMO biocatalyst was carried out either in an LH210 2 litre stirred tank fermenter (Bioprocess Engineering Services, Charing, Kent, UK) or an LH2000, 7 litre fermenter (Bioprocess Engineering Services, Charing, Kent, UK). The fermenters were fitted with two (LH210) or three (LH2000) top-driven, equally spaced six bladed turbines and four diametrically opposed baffles. The specification of each fermenter type can be found in Appendix I. pH was monitored by the use of an Ingold pH probe (Ingold, Mettler-Toledo, Leicester, UK), and DOT was measured with a polarographic oxygen electrode (Ingold, Mettler-Toledo, Leicester, UK). The pH was maintained at 7 (± 0.1) via the addition of 3M NaOH and 3M H_3PO_4, and the temperature at 37°C (± 0.5 °C). In addition, the vessels were aerated at a rate of 0.67 vvm with sterile air, via a submerged sparger.

Inlet and exhaust gasses were filtered using 0.2µm filters (Gelman Sciences, Ann Arbor, MI, USA) and the composition of exhaust gas determined using a mass spectrometer (Prima 600, VG Gas Analysis, Winsford, Cheshire, UK). On-line data logging of fermenter operating variables, e.g. impeller speed, DOT, temperature, pH were recorded with an RT-DAS program (real-time data acquisition system) (Acquisition Systems, Guildford, Surrey, UK) along with the exhaust gas measurements.

3.1.1.3.1. LH210 Fermenter

The LH210 fermenter containing 1.5 L of complex growth medium and the fermenter auxiliaries were sterilised by autoclaving at 121 °C for 20 minutes prior to the addition of filter sterilised ampicillin (100 $mg.L^{-1}$). The fermenter was then inoculated with 100ml of *E.coli* Top10 [pQR239] culture, grown overnight in a shake flask.

Initially the impeller speed during the batch fermentation was set at a rate of 800 rpm. This rate was increased incrementally to a maximum of 1200 rpm to maintain the DOT above zero until the culture biomass concentration had reached 4 g $dcw.L^{-1}$, when expression of CHMO was induced. Addition of L-arabinose solution to the fermenter via a 0.2μm filter to a final concentration of 0.15 % w/v induced the heterologous expression of CHMO *E.coli* Top10 [pQR239]. In order that the DOT did not rise above 5% (Doig *et al.*, 2001) during induction, the impeller speed was reduced when necessary. Three hours after induction, the cells were harvested from the fermenter.

3.1.1.3.2. LH2000 series I Fermenter

Fermentations were also performed in a 7 litre (5.3 litre working volume) LH 2000 series I fermenter. The fermenter containing 5.0 litres complex medium was steam sterilised at 121°C for 20 min. Filter sterilised ampicillin was added to the medium just prior to inoculation with 300 mL of shake flask cell culture .

During the growth phase, the DOT was maintained above 20%. This was achieved by increasing the agitation speed, (set initially at 500 to 600 rpm) each time the DOT dropped below 30% to a final rate of 1200 rpm or until expression of CHMO was induced. Induction was achieved by adding 0.15 % w/v L-arabinose once the biomass reached 4 gl^{-1}. Cells were harvested 3 hr after induction. The DOT was close to 0% for the entire induction period.

3.1.1.4. Biocatalyst Storage

The cells harvested from the 2 L or 7 L fermenters were analysed to quantify the level of cellular CHMO activity (section 3.1.3.3). The cells were then either used immediately in the same medium to carry out a biotransformation or where appropriate stored in a fridge at 4°C. The maximum period of time whole cell biocatalyst broth to be used for biotransformation was stored in the fridge was 7 days.

3.1.2. Whole Cell Biotransformation

3.1.2.1. Small-scale BVMO Biotransformation

Small scale biotransformations were carried out in 1.0 L baffled shake flasks to assay the biocatalytic activity of the whole cell *E.coli* Top10 [pQR239] biocatalyst under various process conditions.

Fermentation broth from whole-cell CHMO production (section 3.1.1) was centrifuged at 4000 rpm for 20 minutes at 4°C (Heraeus Megafuge, Heraeus Instruments Ltd, Brentwood, Essex, UK). Cells were re-suspended in 50 mM phosphate buffer, pH 7 (unless stated otherwise) to a final biomass concentration of $1gl^{-1}$ and volume of 20ml. Glycerol was then added to a final concentration of 10 gl^{-1}. In a shake flask the biotransformation components were incubated in an orbital shaker at 37°C and 200 rpm for 10 minutes.

Addition of neat ketone, to a final concentration of 1.0 gl^{-1} started the biotransformation, which was carried out over a 1-hour period in an orbital shaker at 37°C and 200 rpm. Samples (1ml) of the biotransformation were withdrawn at 10-minute intervals and analysed by GC as described in section 3.1.3.5.

3.1.2.2. Lab-scale (2L) BVMO Biotransformation

Biotransformations were carried out in a LH210 series stirred tank reactor (STR), with a working volume of 1.5 L (Bioprocess Engineering Services, Charing, Kent, UK). The STR was fitted with two equally spaced six bladed

turbines, which were top driven and four diametrically opposed baffles. The reactor comprised a glass vessel with an aspect ratio of 1.82. The ratio of the vessel diameter to the impeller diameter was 2.4. The pH was measured by an Ingold pH probe (Ingold, Mettler-Toledo, Leicester, UK), and maintained at 7 (± 0.1) via the addition 3M NaOH and 3M H_3PO_4. DOT was measured with a polarographic oxygen electrode (Ingold, Mettler-Toledo, Leicester, UK). The vessel was aerated with 1 vvm air via a submerged sparger, and agitation was controlled to maintain DOT above zero by setting the impeller speed at 800 rpm in all biotransformation experiments. The temperature was maintained at 37 °C and antifoam (polypropylene glycol 2000) was added manually as necessary.

The biocatalyst *E.coli* Top10 [pQR239] produced by fermentation (Section 3.1.1.3) was transferred, typically suspended in the spent complex growth medium to the STR. In some instances, an alternative biotransformation medium was used; Table 3.0 lists these media types. The use of an alternative medium type required the centrifugation of the biocatalyst fermentation broth, by batch centrifugation at 4000 rpm for 20 minutes at 4°C (Heraeus Megafuge, Heraeus Instruments Ltd, Brentwood, Essex, UK). The wet cell paste was then resuspended in an alternative medium to the required biocatalyst concentration. In one experiment the cell paste was washed by re-suspension in an equal volume of phosphate buffer, a second centrifugation step was then undertaken prior to cell resuspension.

All biotransformations were run at 37 °C, typically for a period of 3 hours. The ketone substrate was fed at varying flow rates (0.5 to 2.0 $gl^{-1}.h^{-1}$) to the STR via a peristaltic pump (Model 205U/BA, Watson Marlow, Falmouth, Cornwall, UK) from an external reservoir. Approximately every 10 minutes the biotransformation was sampled by withdrawing up to 6 ml of reaction broth. Analysis of samples was carried out as necessary (section 3.1.3 and section 4.1.1).

Media Types	Constituents
Complex medium	10 gl^{-1} of glycerol, 10 gl^{-1} of NaCl, 13 gl^{-1} of Tryptone Peptone, 27 gl^{-1} of yeast extract.
Phosphate buffer	50 mM phosphate buffer
Defined medium	4.25 gl^{-1} KH_2PO_4, 3.15 gl^{-1} $NaHPO_4.12H_2O$, 9 gl^{-1} $NaNO_3$, 0.1 gl^{-1} $MgSO_4$, 0.05 gl^{-1} $CaCl_2.2H_2O$, 0.024 gl^{-1} $FeSO_4.7H_2O$, 0.009 gl^{-1} $MnSO_4.H_2O$, 0.006 gl^{-1} $ZnSO_4.7H_2O$, 0.004 gl^{-1} $CoCl_2.6H_2O$.
Water medium	RO water

Table 3.0: Biotransformation medium. Glycerol, a reductant for the biotransformation was added to each medium preparation to a final concentration of 10 gl^{-1} glycerol.

3.1.3. Analytical Methods

3.1.3.1. Biomass Quantification

Growth of *E.coli* Top10 [pQR239] in the STR's was profiled by measuring changes in the optical density (OD) of withdrawn samples of culture medium. OD measurements were taken at 670 nm (OD_{670}) using an Uvikon 922 variable wavelength spectrophotometer (Kontron, Watford, Herts, UK). Where necessary the samples were diluted appropriately with RO water so that the OD measurements were between 0.1-1.0 absorbance units. The standard error (Appendix II) for OD_{670} measurements was determined from 10 duplicate measurements (biomass concentration range 0.1 – 10 g dcw.L^{-1}) of a single batch fermentation and found to be 0.01 absorbance units.

To determine the dry cell weight (dcw) biomass concentration of a fermentation sample, pre-dried and pre-weighed duplicate 1.5 ml Eppendorf tubes were filled with a single 1 ml aliquot of the cell culture sample. Samples were then centrifuged at 13,000 rpm for 2 minutes, after which the supernatant was discarded. The remaining wet-cell pellet was dried by placing the sample in an oven at 100 °C for 24 hours. Dried samples could then be weighed to calculate the dcw. The standard error (Appendix II) for dcw measurement was determined from duplicate measurements of each of 10 cell culture samples (biomass concentration range 5 – 9 g dcw.L^{-1}) from a single batch fermentation and found to be 0.37 g dcw/L.

A correlation between OD_{670} and dcw was determined from the OD_{670} and dcw measurement of 5 replicate cell culture samples (average concentration 18.8 g dcw.L^{-1}). One g dcw.L^{-1} of *E.coli* Top10 [pQR239] cell culture in complex growth medium was determined to equal 1.83 absorbance units.

3.1.3.2. Protein Assay

The protein content of biotransformation samples was determined by means of a Bradford Assay (Bradford, 1976). Biotransformation samples were clarified by centrifugation at 13,000 rpm for 2 minutes (Heraeus Biofuge 13).

20 µl of supernatant was then mixed by vortexing with 1 ml of a 1 to 5 RO water dilution of Coomassie blue G-250 assay reagent (Biorad Hemel Hempstead, Herts, UK). Samples were transferred (after being left to stand for 5 minutes) to 1ml cuvettes and the OD at 595 nm measured. All measurements were taken on an Unicam UV/Vis spectrophotometer (Spectronic Analytical Instruments, Leeds, UK) after adjusting the absorbance to zero with a sample of RO water. Where necessary the biotransformation samples were diluted appropriately with RO water so that the OD measurements were between 0 - 1.0 absorbance units.

A calibration curve (Appendix III) linking the OD_{595} to protein concentration was constructed using a series of standard samples of known protein concentration. Standard samples had bovine serum albumin protein concentration in the range of 0.2 to 10 $mg.ml^{-1}$; prepared in RO water. The standard samples were then assayed as for the clarified biotransformation samples. This enabled the prediction of biotransformation samples where the protein concentration was unknown.

3.1.3.3. CHMO Intracellular Activity Assay

The intracellular activity of CHMO in *E.coli* Top10 [pQR239] was monitored using an NADPH rate assay, as the oxidising power of the co-factor NADP+ is used in stoichiometric amounts by the reductant CHMO during the formation of the combined lactone from the ketone (Figure 2.1). This was undertaken in a 30°C temperature controlled variable wavelength Kontron Uvikon 922 spectrophotometer (Watford, Herts, UK). Oxidation of NADPH (ε = 6.22 ml/µmol/cm) was monitored over time at a wavelength of 340nm. The method is described by (Donoghue *et al.*, 1976).

CHMO was prepared from *E.coli* Top10 [pQR239] culture broth sampled from a fermentation or biotransformation experiment. Medium supernatant was removed after micro-centrifugation at 13,000 rpm for 1 min (Heraeus Biofuge 13, Herts, UK). The cell pellet was then re-suspended in a solution of Tris-HCL buffer (50 mM, pH 9.0), and bovine serum albumin (5 $mg.ml^{-1}$). To release the CHMO enzyme, cell disruption by sonication (Soniprep 150;

Sanyo, Crawley, UK) was undertaken in an ice-bath at a wavelength of 8 μm for 5 cycles of 10 seconds, with 10 seconds between cycles. Finally, clarification was by micro-centrifugation at 13,000 rpm (10,000 g) for 2 minutes. The supernatant was then used neat or diluted in RO water as the CHMO source.

In a final volume of 1ml the following reagents were added to a 1.5ml cuvette: Tris-HCL buffer (50 mM, pH 9.0), bovine serum albumin (5 mg/ml), NADPH (0.161 mM), CHMO (Concentration was such that rate of activity was initially linear). The background rate of NADPH activity monitored at 340 nm (A_{340}) was measured over a period of 5 minutes. A second rate of NADPH activity was then measured for a further 5 minutes, on addition of cyclohexanone (2mM) to the cuvette.

The rate of NADPH oxidation could then be calculated by subtracting the background rate from the second NADPH oxidation rate. The activity of CHMO (Ca) was expressed in international units (U) with one unit being defined as the number of μmoles of NADPH consumed per minute. The calculation of the concentration of NADPH was based on Beer's Law (Section 1.2.9) and so to determine the number of units of CHMO activity, equation 3.1 was used. The error associated with this assay was ±150 U based on the mean standard deviation (Appendix IV).

$$Ca = \frac{\Delta A_{340}}{I \cdot \varepsilon \cdot \Delta t}$$

ε =extinction coefficient ($\mu mol.ml^{-1}.cm^{-1}$), CHMO ε = 6.22 $\mu mol.ml^{-1}.cm^{-1}$

I = light path length (cm)

Equation 3.1

3.1.3.4. Glycerol Assay

An enzymatic assay (Boehringer Mannheim Roche, Roche-Biopharm GmbH, Darmstadt, Germany) was used for the determination of glycerol in fermentation and biotransformation samples.

Glycerol was phosphorylated by adenosine-5'-triphosphate (ATP) to L-glycerol-3-phosphate in the presence of glycerokinase (GK), as shown by equation 3.2.

$$Glycerol + ATP \xrightarrow[GK]{} L-glycerol-3-phosphate + ADP$$

Equation 3.2

The adenosine-5'-diphosphate (ADP) formed in the above reaction was reconverted by phosphoenolpyruvate (PEP) into ATP, a reaction catalysed by pyruvate kinase (PK) and resulting in the formation of pyruvate (equation 3.3)

$$ADP + PEP \xrightarrow[PK]{} ATP + pyruvate$$

Equation 3.3

In the presence of the enzyme L-lactate dehydrogenase (L-LDH), pyruvate is reduced to L-lactate by reaction with reduced nicotinamide-adenine dinucleotide (NADH); as shown in equation 3.4.

$$Pyruvate + NADH + H^+ \longrightarrow (L)lactate + NAD^+$$

Equation 3.4

The amount of NADH oxidised in the above reaction is stoichiometric to the amount of glycerol. NADH is determined by measurement of light absorbance at 340 nm. The error associated with this assay was ± 0.07 gl^{-1} based on 6 independently prepared standard solutions (concentration range 4 to 9 gl^{-1})

3.1.3.5. Ketone and Lactone Quantification

The separation and quantification of ketone and lactone from ethyl acetate preparations was performed on a Perkin-Elmer autosystem XL-2 gas chromatograph (Perkin-Elmer, Norwalk, CT, USA), fitted with an Alltech series AT-1701 (Alltech, Carnforth, Lancs, UK) and linked to an FID detector. Data collection and reduction was performed using PE Nelson Turbochrom™ software (PerkinElmer, Wellesley, MA USA).

3.1.3.5.1. Sample Preparation and GC Operation

To prepare aqueous samples from biotransformation experiments for GC analysis, a 1 ml sample volume was centrifuged at 13,000 rpm for 4 minutes (Heraeus Biofuge 13). A 0.5 ml sample of supernatant was transferred to a fresh vial before addition of an equal volume of ethyl acetate. The two-phase mixture was then vortex mixed for 10 seconds in a 2.2 ml Eppendorf tube. A 10 second vortex period was determined to be enough for the partioning of analytes between phases to reach equilibrium (data not shown). The mixture was centrifuged for a further 2 minutes at 13,000 rpm, the organic layer removed and if necessary diluted with ethyl acetate before analysis by GC. Analyte ethyl acetate samples prepared for calibration of GC chromatographs were transferred directly to the GC auto-sampler. The conditions used to operate the gas chromatography (GC) apparatus are given in table 3.1.

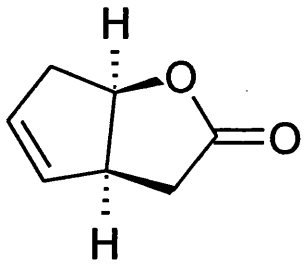
Operational parameter	Configuration
Column	Reverse phase silica, (30 m x 0.54 mm)
Injection volume	1 µl
Carrier gas flow rate	4.0 ml.min^{-1}
Column temperature program	100 °C held for 5 minutes then raised by 10 °C/min up to a maximum temperature of 240 °C.
FID temperature	250 °C

Table 3.1: Gas chromatography operating conditions.

3.1.3.5.2. Quantification of the GC Chromatographs

Unknown analyte concentrations in the ethyl acetate phase of samples were calculated using predetermined calibration curves. These curves were based on independently prepared solutions (concentration range 0.1 gl^{-1} to 2.5 gl^{-1}) of ketone and lactone[1] (lactone[1] is depicted in Figure 3.0) in ethyl acetate. Typical calibration curves of analyte concentration against FID response are shown in Appendix V. Lactone[1] is commercially available, but the combined lactones; (-) 1(S), 5(R)2-oxabicyclo(3.3.0)oct-6-en-3one and (-) 1(R), 5(S) 3-oxabicyclo(3.3.0)oct-6-en-2-one, produced during the model biotransformation are not. The GC response time for lactone[1] was the same as for the combined lactones (Appendix VI), it was therefore considered suitable to produce standards for combined lactone calibration. Lactone[1] is the opposite enantiomer of (-) 1(S), 5(R)2-oxabicyclo(3.3.0)oct-6-en-3one.

The standard error of GC analysis was determined from 10 duplicate sample measurements (concentration range 0.1-2.5 gl^{-1}) and found to be 91027 u/V for ketone and 156025 u/V for lactone[1]. This equates to a standard error of 0.04 and 0.07 gl^{-1} for ketone and lactone respectively.



4

Figure 3.0: Lactone[1] 1(R), 5(S) 2-oxobicyclo(3.3.0)oct-6en-3one.

3.1.3.5.3. Ethyl Acetate Partition Coefficient

Quantification of the ketone and lactone concentrations in biotransformation samples required the equilibrium water-ethyl acetate partition coefficient to be determined for these compounds. Aqueous solutions of the ketone and lactone[1] were made up at various concentrations between 0.1 – 1.0 gl^{-1}. 10ml of each solution was contacted with 10ml of ethyl-acetate at room temperature with agitation at 200rpm for 24 hours in sealed glass vials. The solvent phase was analysed for ketone and lactone concentration using the gas chromatography method. The ethyl-acetate/ water partition coefficient, K, was subsequently determined as the ratio between the equilibrium concentration of solute in ethyl-acetate, C_e, and the equilibrium concentration of solute remaining in the aqueous phase, C_{aq} using equation 3.5. It was determined that 100% of the ketone and 88% of the lactone[1] was present in the solvent phase at equilibrium. The concentration of lactone[1] remaining in the aqueous was determined by mass balance. The partition coefficient for ketone and lactone[1] was therefore 1.0 (nearing infinity) and 7.11 ±0.12 respectively.

$$K = \frac{C_e}{C_{aq}}$$

Equation 3.5

3.2. Results

3.2.1. Biocatalyst Production

Whole-cell catalysed Baeyer-Villiger type oxidation of cyclic ketones yielding chiral lactones require the host to be cultivated and CHMO expressed. *E.coli* TOP10 [pQR239] expressing CHMO from *Acinetobacter calcoaceticus* NCIMB 9871, was cultivated at 2 and 7 litre scale in batch fermentations. The cloning and fermentation characterisation of *E. coli* TOP10 [pQR239] have previously been reported (Doig *et al.*, 2001) as described in section 1.1.

Production of the *E. coli* TOP10 [pQR239] biocatalyst by batch fermentation is described in this section. The aim of the work was to provide a protocol for the production of the biocatalyst and to establish a reproducible and characterised Baeyer-Villiger bioconversion.

3.2.1.1. Growth Characteristics

The production protocols for *E.coli* TOP10 fermentation at 2 L and 7 L scales were developed to be similar in order that inter-batch variation was minimised. For instance, the same complex growth medium was used throughout this research. The operation of the reactors was based on previously published work (Doig *et al.*, 2001) and sought to maximise both the final biomass concentration and specific intracellular CHMO titre.

A typical 2 L fermentation profile is depicted in Figure 3.1. The growth curve is characterised by a short lag phase, which is due in part to the seed medium and fermentation medium being identical. This can also be explained by the relatively high volume of the seed culture, which was 6% v/v of the working reactor volume. The final biomass concentration of 8.5 gdcw.l^{-1} was achieved with a maximum specific growth rate of 0.77 h^{-1}. The average final biomass concentration for 2 L fermentations was 7.6 ± 1.1 g dcw.l^{-1}. The oxygen uptake rate (OUR) and carbon dioxide evolution rate (CER) was controlled initially such that it was rapidly increasing. This control was achieved by maintaining the dissolved oxygen tension (DOT) above 10% so

that oxygen supply was not limiting growth. A step-wise increase in the agitation speed from 600 rpm provided a means of controlling the DOT above 10%. DOT levels dropped during the late log-phase of culture growth. This was due to the stirrer speed been restricted to a maximum of 1250 rpm and the on-set of CHMO expression by induction of the L-arabinose controlled promoter. On induction of the CHMO expression by *E.coli* Top10 [pQR239] at 225 min, the fermentation was controlled to achieve a DOT <5%, whilst maintaining a high OUR since previous work (Doig *et al.*, 2001) had indicated that this approach gave the highest titre of intracellular CHMO. An OUR of approximately 110 $mmol l^{-1} hr^{-1}$ was maintained for 1.5 hours post induction, then the OUR dropped to below 60 $mmol l^{-1} hr^{-1}$ as the cells entered the stationary phase of growth. At this point the DOT started to increase. The fermentation was stopped and the biocatalyst harvested from the fermenter.

A similar profile exists for the 7 L fermentation (Figure 3.2) since the objective was simply to produce the same titre of CHMO at a larger scale. As for the 2 L fermentation the growth curve is characterised by a short lag phase since the fermentation and seed medium are identical. The seed volume was 5.7% v/v of the working volume of the fermenter. The average final biomass concentration for 7 L fermentations was higher than the 2 L at 12 ± 3 g $dcw.l^{-1}$. In the example shown the biomass concentration had reached 11 gl^{-1} with a maxium specific growth rate of 0.78 h^{-1} when the fermentation was stopped. The OUR and CER was high in the 90 to 110 $mmol.l^{-1}.h^{-1}$ region and the DOT was recording 0% at the end of the fermentation indicating that the culture was yet to enter the stationary phase of growth in the example shown. The operating method for the production of CHMO at a 7 L scale (section 3.1.1.3.2) is similar to the 2 L scale in that termination of the fermentation is timed at 3 hours post induction. It is recognised that the yield of biocatalyst and titre of CHMO are not optimised for the fermentation method described, although sufficient yields were achieved for the purpose of this study.

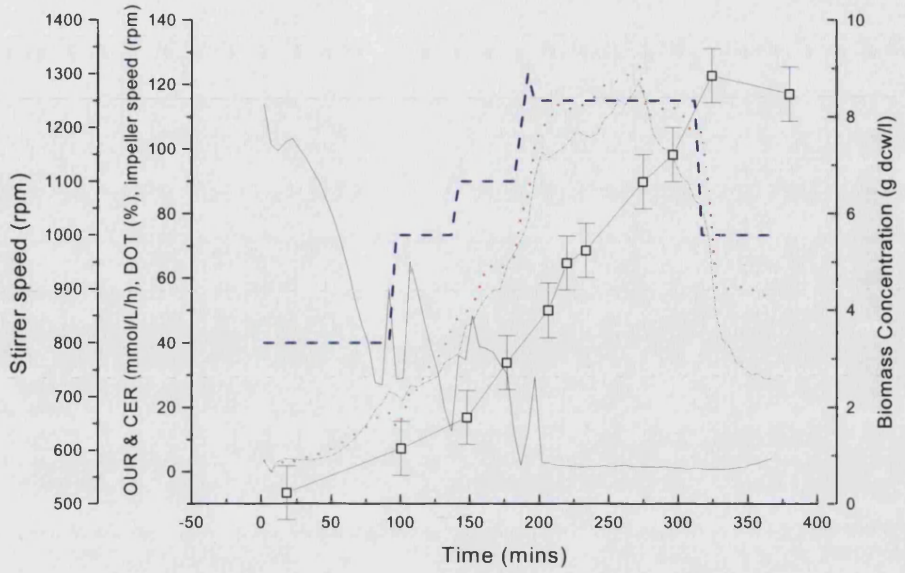


Figure 3.1: Representive 2 L fermentation profile for *Escherichia coli* Top10 [pQR239]. (□) Biomass concentration. (---) Impeller speed. (—) Dissolved oxygen tension (DOT). (·····) Oxygen uptake rate (OUR). (– - –) carbon evolution rate (CER).

The DOT profile in Figure 3.2 falls as OUR increases with increasing biomass concentrations. As the DOT falls below 10% at 150 minutes the stirrer-speed increases exponentially to a maximum of 1200 rpm. This operation was controlled by a programmable logic controller set-up to increase the stirrer speed and hence OUR when DOT fell below 10%. Fermentations were always operated not to limit oxygen to the growing cells up to induction of CHMO expression (180 minutes for figure 3.2), after which the level of DOT was maintained at a recorded 0% i.e. all oxygen inputted in to the system was being dissolved in to the liquid phase and consumed by the cells. The scale-up rationale for the 2 L to 7 L fermentation is discussed in section 3.2.3.

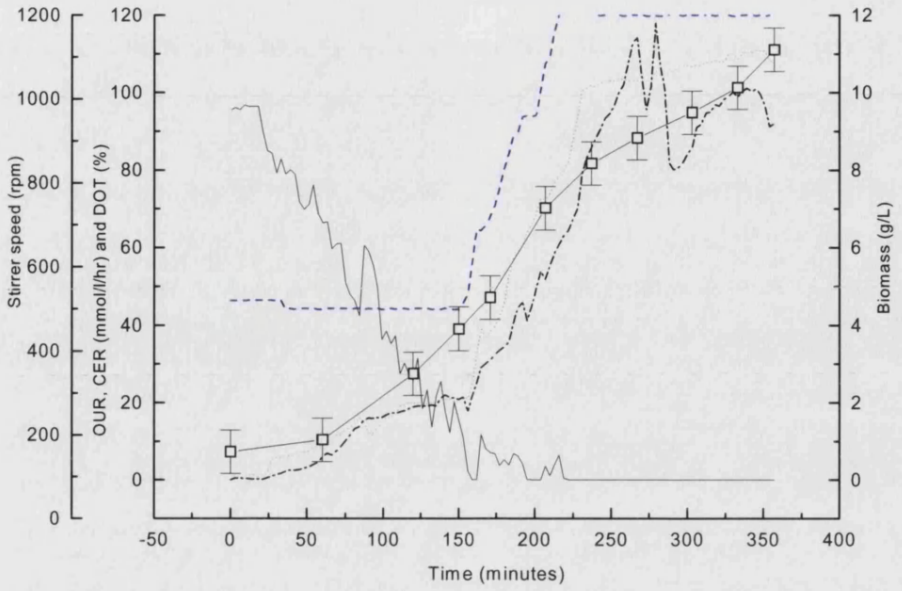


Figure 3.2 Representative 7 L fermentation profile for *Escherichia* coli Top10 [pQR239]. (□) Biomass concentration. (---) Impeller speed. (—) Dissolved oxygen tension, (DOT), (·····) Oxygen uptake rate (OUR), (- –) carbon evolution rate (CER).

3.2.1.2. Medium Protein Content

The volumetric scale at which the CHMO expressing *E.coli* biocatalyst was produced affected the relative concentration of cells contained in the fermentation broth. Both fermentation methods used the same complex medium (section 3.1.1). Conventionally, the whole-cell biotransformation of ketone to combined lactone has been undertaken in the spent complex fermentation broth (Doig *et al.*, 2002). Variation in the constituents of the spent broth between inter-fermentation batches resulting because of the fermentation process will be carried through to the next step in the bioprocess, the biotransformation step. To qualify variation between biotransformation medium the extra-cellular protein concentration of samples was analysed using the Bradford assay.

The scale at which the *E.coli* cells were grown resulted in different extra-cellular protein concentration profiles within clarified biotransformation broth samples (Figure 3.3). It is apparent by the example shown in Figure 3.3 that the broth from 7 L fermentations contains more than twice the concentration of extra-cellular protein and that this trend is observed throughout the period of the biotransformation. In both cases the relative extra-cellular protein concentration increases during the biotransformation.

The whole-cell biotransformation of bicycloheptanone using spent complex broth from the 7L fermenter resulted in increased levels of extra-cellular protein compared to the 1.5L scale fermenter broth (Figure 3.4). When the biocatalyst was separated from the spent fermentation broth and re-suspended in alternative biotransformation medium (Figure 3.4) a significant reduction in the concentration of protein was observed. Such variation between biotransformation media types is carried over to subsequent biotransformation steps and will contribute to complexities during the analysis of operations using near infrared spectroscopy.

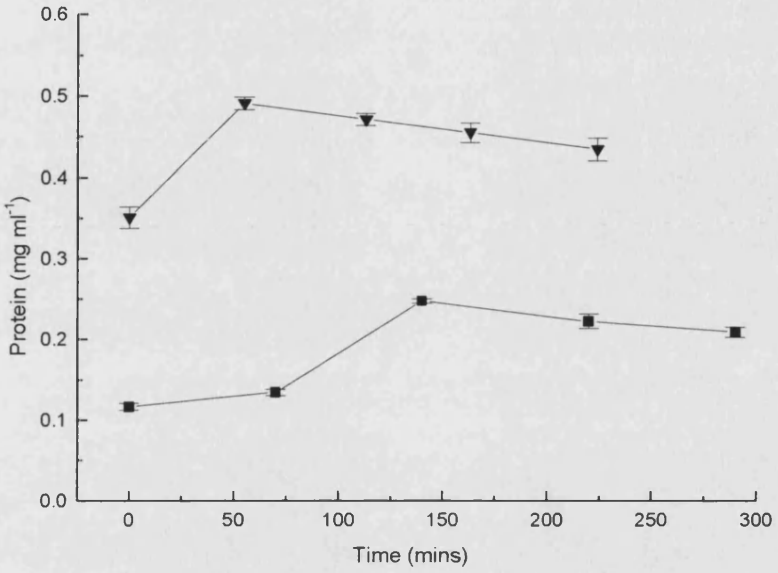


Figure 3.3: Fermenter configuration and protein content of medium during biotransformations. (■) Representative 2 L spent fermentation broth (▼) Representative 7 L spent fermentation broth. Error-bars are based on the standard deviation of 3 replicates for each sample.

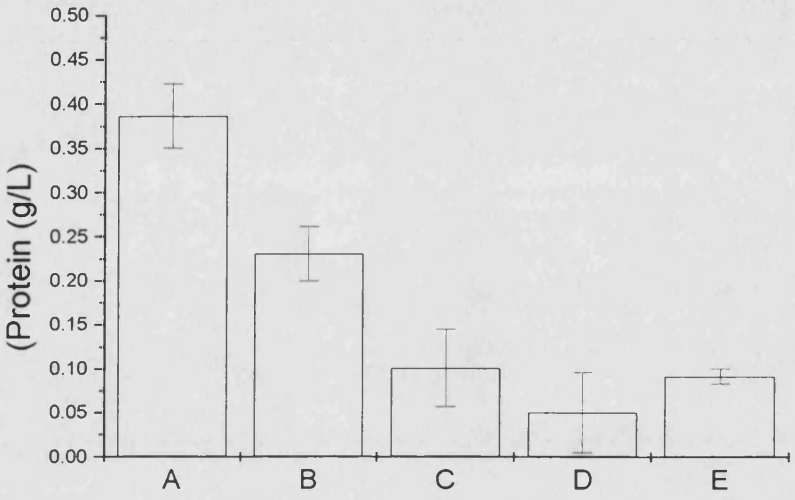


Figure 3.4: Protein content of biotransformation medium. Results show the average and standard deviation for 5 samples taken at the start of individual biotransformations. (A) = spent medium from a 7 L fermentation; (B) = spent medium from a 2 L fermentation; (C) biocatalyst resuspended in a minimal defined medium; (D) = biocatalyst resuspended in phosphate buffer; (E) biocatalyst resuspended in water.

3.2.1.3. Biocatalyst Activity

The expression of CHMO was induced by the addition of 0.15 % w/v L-arabinose to the culture broth. This induction occurred when the optical density of withdrawn clarified broth sample reached 9 to 10 absorbance units (approximately 5 g dcw.l^{-1} biomass) at a wavelength of 670 nm. An optical density of 9 to 10 represents the late exponential growth phase for the operated batch conditions. The results shown in Figure 3.5 show that the intracellular CHMO titre of the *E.coli* biocatalyst increased sharply from the point of induction with undetectable CHMO activity prior to this. The cells were harvested three hours post-induction, typically as the culture entered the stationary growth phase. At the point of harvesting the intracellular CHMO titre reached a maximum value, which was 2290 UI^{-1} for the fermentation data shown in Figure 3.5. The average final CHMO titre for 2 L and 7 L fermentations was 2100 ± 380 UI^{-1} and 3200 ± 600 UI^{-1} respectively. The maximum specific intracellular titre for 2 L and 7 L fermentations was 228 ± 38 U/g and 254 ± 56 U/g respectively.

The whole-cells were harvested from the fermenter in the spent fermentation broth. Post harvesting the biocatalyst was either stored suspended in the broth in a refrigerator (4°C) or transferred to the biotransformation reactor. Storage of the biocatalyst in a refrigerator was possible for up to 7 days (approximately 170 hours) without an observable loss in intracellular activity (Figure 3.6). Figure 3.6 shows that after a storage time of 170 hours, the intracellular activity of the biocatalyst starts to fall. Biocatalyst was routinely stored under refrigerated conditions for up to 5 days prior to a biotransformation.

The stability of the biocatalyst was observed during a control CHMO catalysed Baeyer-Villiger oxidation. The experiment was carried out as described in Section 3.1.2.2. No ketone was fed into the reactor during this control biotransformation. The absence of ketone and therefore of the biotransformation reaction provided samples indicating a measure of the biocatalyst stability during operational conditions. At regular intervals during

the control experiment 100ml samples were withdrawn. The intrinsic specific activity of the *E.*coli biocatalyst samples was assayed over a 1-hour period in a small scale batch biotransformation of the ketone substrate as described in section 3.1.2.1. Figure 3.7 indicates that the normalised intrinsic specific biocatalyst activity was maintained for 4 hours before a significant loss was observed. The fitted curve (Figure 3.7) approximates biocatalyst stability and has the form shown in Equation 3.6.

Equation 3.6

$$Y = -1.8004x10^{-8}t^3 + 2.0172x10^{-6}t^2 + 1.3426x10^{-4}t + 0.99833$$

where y is the residual biocatalyst activity and t is the duration of the reaction.

As indicated, the specific intracellular CHMO titres of *E.coli* biocatalyst batches stored at 4°C were stable for 7 days. Batches stored over a 7 day period were also observed to maintain the intrinsic specific whole cell activity (data not shown). The observed decrease in the intrinsic specific CHMO activity of the biocatalyst after 4 hours in a STR at 37°C and pH 7 was a phenomenon dependent on the operating conditions in the STR.

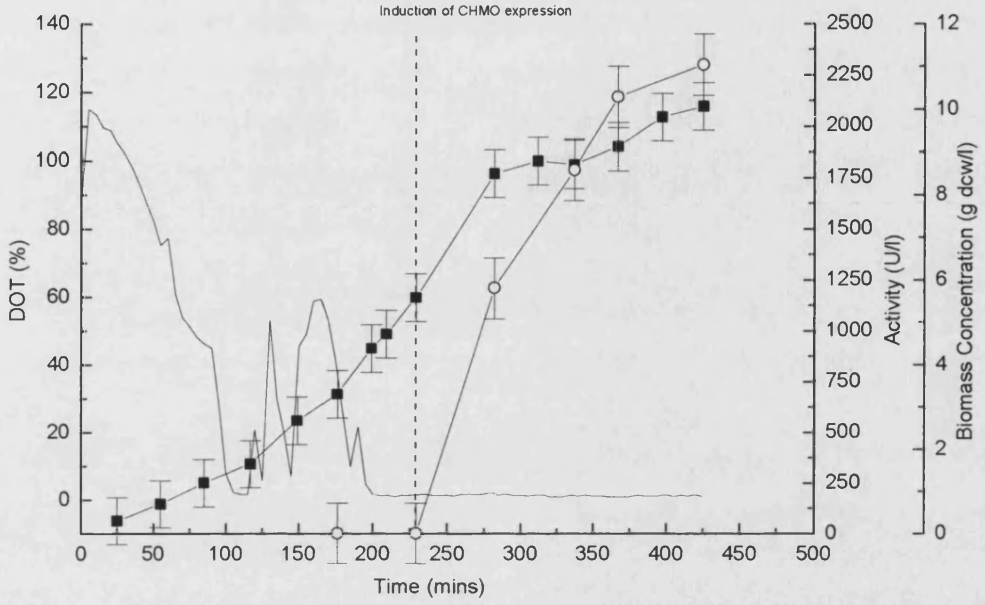


Figure 3.5 CHMO expression during a 2 L *Escherichia* coli Top10 [pQR239] fermentation. (■) Biomass concentration. (—) Dissolved oxygen tension, (DOT). (○) Intrinsic CHMO activity.

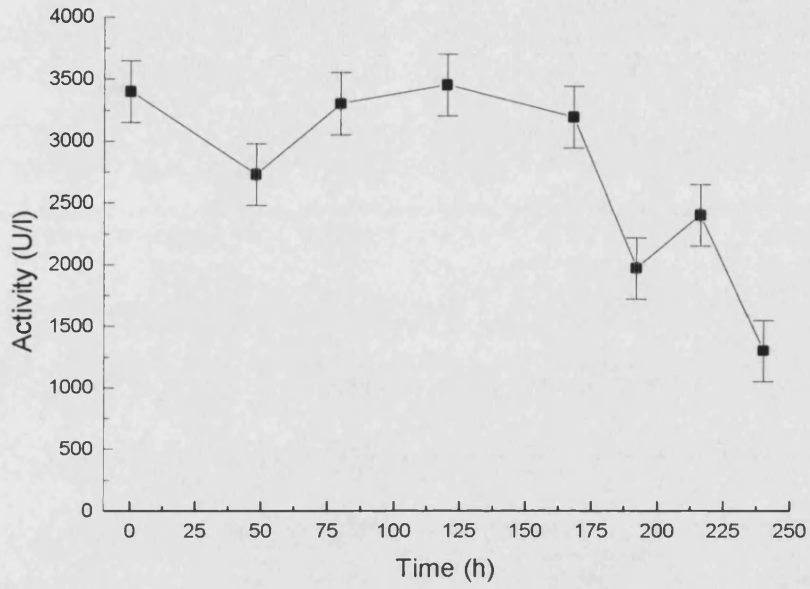


Figure 3.6 Intracellular CHMO activity of *Escherichia* coli Top10 [pQR239] biocatalyst activity during storage post harvesting. The biocatalyst was stored suspended in the spent fermentation broth in a refrigerator at 4°C.

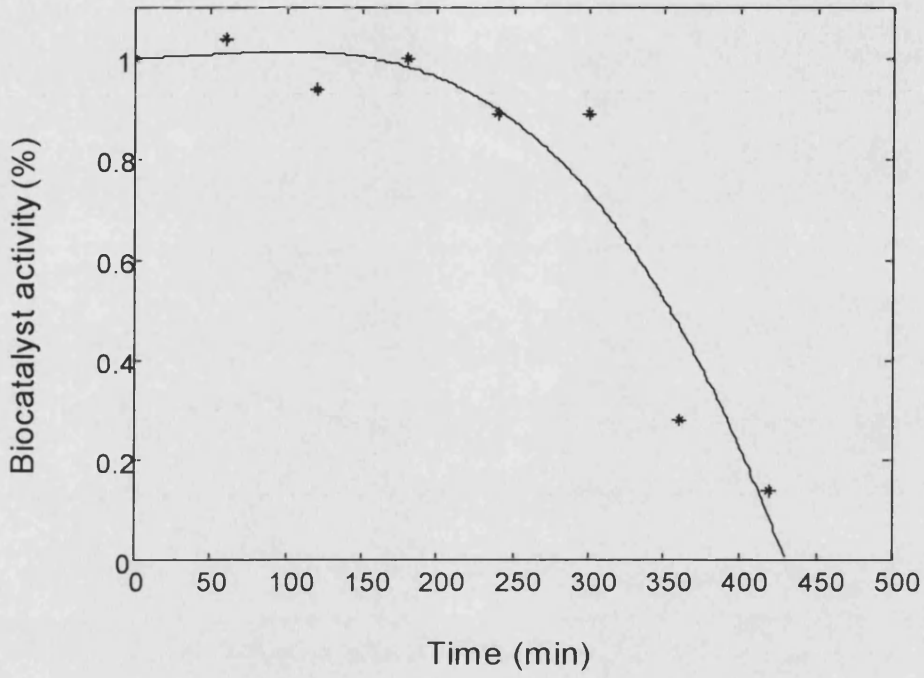


Figure 3.7: Modelled biocatalyst stability. (*) Experimental data derived from a 2 L stirred tank under process conditions (—) fitted curve.(Doig *et al.*, 2003a).

3.2.2. Whole Cell Biotransformation

3.2.2.1. Glycerol Required for Biotransformation

CHMO is a BVMO which uses a flavin coenzyme as an organic cofactor and NAD(P)H as the reductant (Kelly, 1998, Willets, 1997) to carry out the Baeyer-Villiger type oxidation. The *E.coli* TOP10 [pQR239] fermentations were carried out using a Glycerol based growth medium. The glycerol constituted the carbon source during the growth of the culture and the sources of reducing equivalents for whole-cell NAD(P)H cofactor recycle during the biotransformation step. It was important to ensure sufficient glycerol was available to the cells since it was shown (Figure 3.8) that at concentrations of <3 gl^{-1} the initial specific activity of the cell (method is described in section 3.1.2.1) was significantly reduced. The requirement to add sufficient glycerol to the biotransformation medium prior to commencement of the biotransformation was included in the process method. Where the biotransformation medium type was the spent broth of the preceding fermentation, glycerol carried over to the biotransformation step enabled the operation to proceed as expected (Figure 3.9). In contrast, re-suspension of the biocatalyst in defined medium requires the presence of glycerol for NAD(P)H cofactor recycle and without this no reaction is observed (Figure 3.9). All routine biotransformations including biotransformations carried out in broth type medium received a batch addition of 10 gl^{-1} glycerol (Figure 3.10).

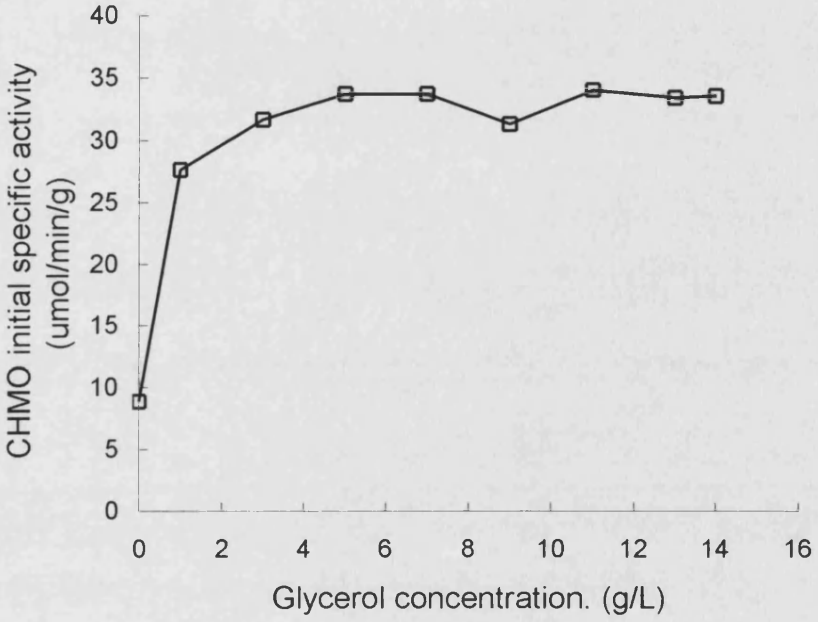


Figure 3.8 : The initial specific activity of whole-cell *E.coli* TOP10 [pQR239] at various concentrations of glycerol. Shake flask biotransformations of ketone to combined lactone. CHMO activity is dependent on reaction mixture glycerol concentration.

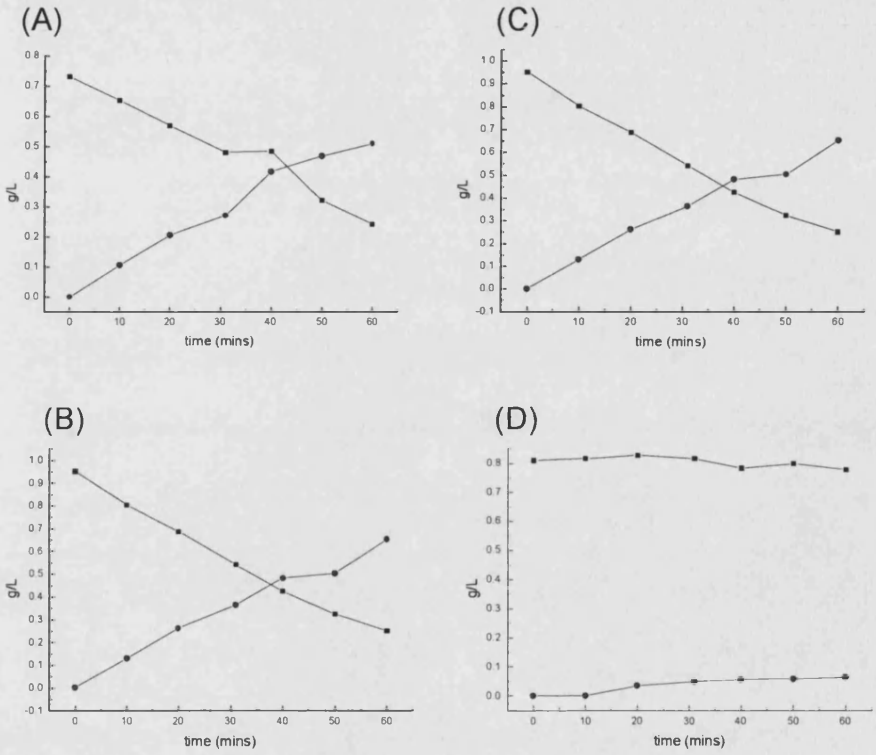


Figure 3.9 Shake flask biotransformations using whole-cell *E.coli* TOP10 [pQR239]. (A) Biotransformation in the spent fermentation broth; Glycerol (10 gl^{-1}). (B) Biotransformation in defined medium, glycerol (10 gl^{-1}). (C) Biotransformation in the spent fermentation broth; no glycerol added. (D) Biotransformation in defined medium, no glycerol added. (■) Ketone, (●) lactone

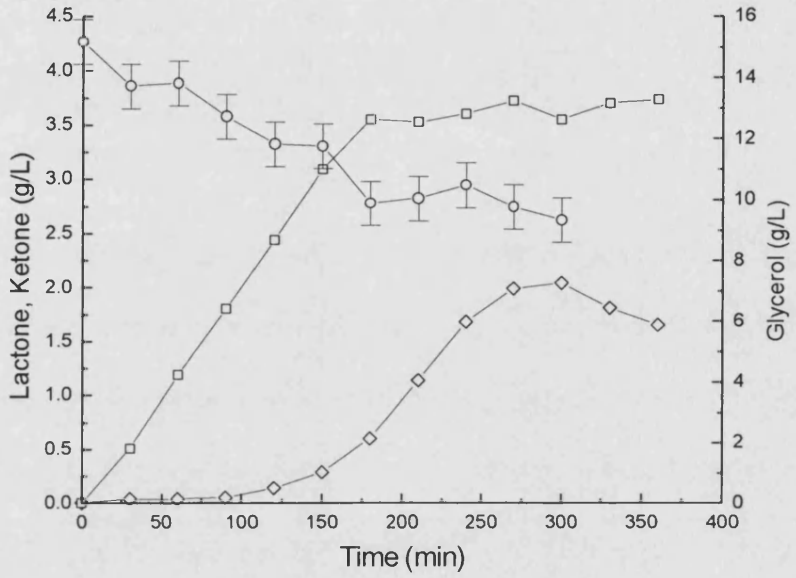


Figure 3.10: Glycerol profile during a representative 2 L *E.coli* TOP10 [pQR239] biotransformation. (□) Lactone. (◇) ketone. (O) Glycerol. Glycerol is added in batch, the ketone substrate is added fed-batch.

3.2.2.2. Biotransformation Profiles

Figure 3.11 shows a range of biotransformation profiles. All fed-batch biotransformations were conducted on a 2L scale in spent fermentation broth as described in section 3.1.2.2. From previous kinetic characterisation it was found that substrate inhibition is a key characteristic of the *E.coli* TOP10 [PQR239] whole cell catalysed oxidation of ketone to its corresponding combined lactones (Doig *et al.*, 2002) as described in section 2.1 of the introduction. In order to address this issue the ketone substrate was fed neat into the reactor to maintain a low ketone concentration in the reaction in an attempt to maximise the specific activity and hence final product concentration attainable. Figure 3.11(A) depicts a representative biotransformation; in this instance ketone was fed at a constant rate of 1.3 $g.l^{-1}.h^{-1}$ into the reactor. Since the ketone feed rate is above the observed CHMO volumetric activity of 1.22 $g.l^{-1}.h^{-1}$, the rate of lactone production is likely limited by:

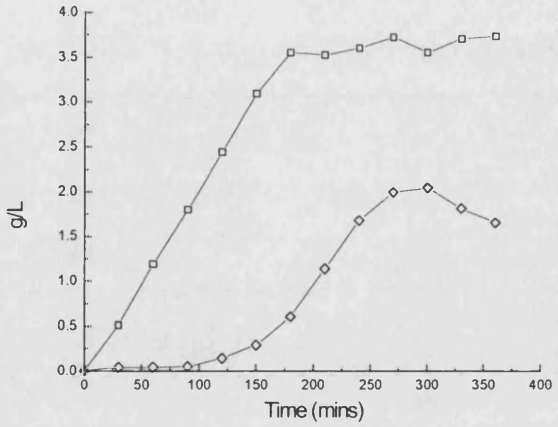
- Lactone inhibition of whole cell CHMO activity (Doig *et al.*, 2003b)
- Time dependent whole cell viability (Chen *et al.*, 2002)
- AND ketone inhibition (Doig *et al.*, 2003b)

On loss of CHMO activity at approximately 160 minutes (Figure 3.11) a dramatic increase in the ketone concentration was observed. By 175 minutes lactone production had peaked. The continual feeding of ketone then resulted in a decrease in the ratio of product formed to substrate added (P/S) (Figure 3.12) up to 300 minutes at which time the ketone feed was stopped. Maximising the yield of P/S is a desirable target for biocatalytic process development. Presence of residual substrate from stereoselective synthesis complicates DSP since substrate and product often exhibit similar purification characteristics. This is the case with ketone, which has been shown to bind competitively to an amberlite XAD4 column along with the lactone products (Alphand *et al.*, 2003). In addition there is a clear benefit to process economics relating to efficient use of raw materials.

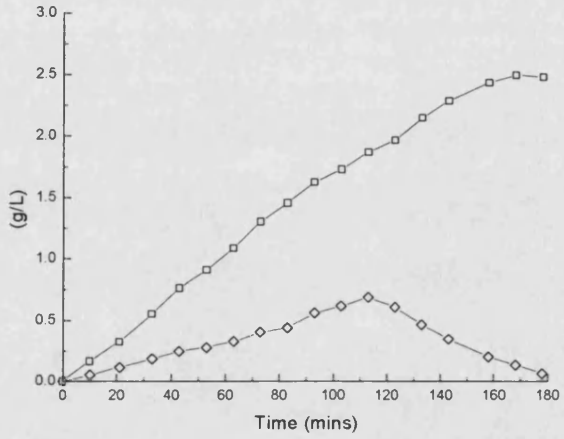
Substrate inhibition can be avoided and a high yield of product on substrate can be achieved by stopping the feeding of ketone to the reactor, as shown in Figure 3.11B. In this instance the substrate feed was stopped at 115 minutes after which the yield (P/S) increased significantly (Figure 3.12). However, the rate of bioconversion slowed prematurely at 180 minutes due to a limiting ketone concentration.

Where the ketone was fed at a rate below the maximum volumetric activity of the biocatalyst (Figure 3.11C) the yield of product on substrate was optimised. The ketone feed rate limits the maximum biocatalyst volumetric activity to 0.71 $g.l^{-1}.h^{-1}$, which is lower than the maximum biocatalyst volumetric activity for reactions A; 1.22 $g.l^{-1}.h^{-1}$ and B; 1.07 $g.l^{-1}.h^{-1}$.

(A)



(B)



(C)

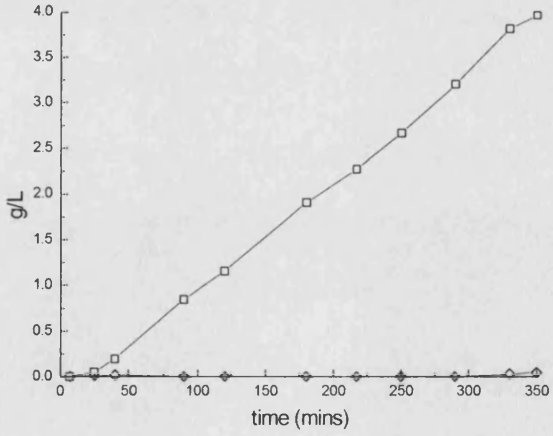


Figure 3.11 (A): Whole-cell *E.coli* TOP10 [pQR239] biotransformation profile with substrate continuosly fed to the reactor at a rate greater than the reaction rate i.e. over-feeding. (B): Biotransformation with substrate feed curtailment. (C): Biotransformation profile with substrate under feeding. (□) Lactone. (◊) ketone.

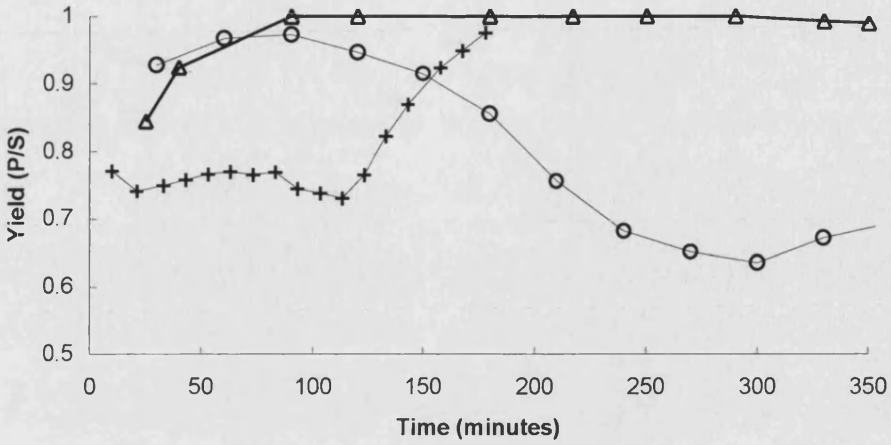


Figure 3.12: The ratio of product formed to substrate added during biotransformation A (O); biotransformation B (+) and biotransformation C (Δ), the profiles of which are depicted in Figure 3.11.

3.2.2.3. Substrate and Product Inhibition

As described in section 2 the biotransformation kinetics for the oxidation of ketone to its corresponding regio-isomeric lactones have previously been investigated (Alphand *et al.*, 2003). The key kinetic characteristics of this strain were found to be severe substrate and product inhibition. This has significant implications on the design of the bioprocess; these are discussed in the discussion section 3.2.3.

3.2.2.4. Biotransformation Medium

The whole-cell Baeyer Villiger biotransformation has typically been undertaken in the spent broth of the preceding fermentation step (Doig *et al.*, 2001). Alternatives to broth medium include water, phosphate buffer, and defined medium. The whole-cell biocatalyst was separated from the broth medium and resuspended in an alternative medium type as described in section 3.1.2. Small-scale biotransformation experiments conducted in 1L

shake flasks at a biocatalyst concentration of 1.25 gl^{-1} indicated that under these conditions the biocatalyst performed equally well in all media types over the period of the biotransformation (Figure 3.13). All media types permitted the Baeyer-Villiger biotransformation to progress, at similar volumetric activities (0.6 $g.l^{-1}.h^{-1}$).

Despite the biocatalyst exhibiting similar activities using the various media types, further work was undertaken at low biocatalyst concentration's at the 2L scale in STR's. The low concentration of *E.coli* Top10 [pQR 239] (0.4 gl^{-1}) resulted in a reduced demand for O_2 such that the OTR rate into the biotransformation medium was not limiting i.e. DOT was not zero (Figure 3.14). The biocatalyst volumetric activity was calculated to be 0.73 $g.l^{-1}.h^{-1}$. A linear increase of lactone was yielded during the time course of the reaction at a rate limited by the ketone feed rate. Ideally, the ketone feed rate would not be limiting, as this would allow for maximum biocatalyst activity during the period of modelled biocatalyst stability. The inability to monitor and control substrate and product concentrations meant that it was difficult to operate the *E.coli* Top10 [pQR239] Baeyer-Villiger process as required. At this low biocatalyst concentration the specific activity of whole-cell CHMO was at least 0.58 $g.l^{-1}.h^{-1}.gdcw^{-1}$ (4.7 $\mu mol.min^{-1}.gdcw^{-1}$) dcw initially and at least 0.14 $g.l^{-1}.h^{-1}.gdcw^{-1}$ (1.1 $\mu mol.min^{-1}.gdcw^{-1}$) on termination of the experiment. The specific activity dropped due to continued growth of *E.*coli Top10 [pQR239]. The growth of *E.coli* (Expression of CHMO by *E.coli* was assumed to be taking place during growth on the spent fermentation medium.) on the spent fermentation broth resulted in a drop in the DOT due to increased cellular oxygen demand until DOT became limiting at approximately 250 minutes.

As indicated in the shake flask experiments (Figure 3.13), the Baeyer-Villiger biotransformation is observed to proceed after re-suspension of the biocatalyst in the phosphate type medium. At the 2L scale operating the bioprocess at a low biocatalyst concentration in water and phosphate medium types (Figure 3.15) the whole-cell *E.coli* culture was not observed to grow. Consequently the DOT wasn't limiting the rate of the biotransformation

reaction at any point during the experiment. The maximum volumetric and specific activity of the whole-cell biocatalyst in phosphate buffer was 0.43 $g.l^{-1}.h^{-1}$ and 0.39 $g.l^{-1}.h^{-1}.gdcw^{-1}$ (3.1 $\mu mol.min^{-1}.gdcw^{-1}$). The maximum biocatalyst activity observed in phosphate buffer was two thirds that observed in the complex medium type. In contrast to the complex medium biotransformation profile (Figure 3.14) the period of maximum activity was maintained for only 175 minutes in the phosphate buffer after which time the volumetric activity dropped. The reduction in biocatalyst activity could have resulted from a loss of whole-cell viability. This is indicated by the rise in the DOT from 25 minutes, which suggests that the culture was not consuming O_2, an important reactant of the Baeyer-Villiger bioconversion.

Re-suspension of the cells in water medium yielded very little lactone and a short period of cell viability as indicated by the DOT (Figure 3.15). The rationale for using a water/ glycerol mix to undertake the biotransformation was that a simple medium type would allow for a controlled biocatalyst concentration and simplified reaction matrix on which to undertake near infrared spectroscopic sampling. The reaction does not proceed in the water medium type because the viability of *E.coli* Top10 [pQR239] is diminished under these conditions. Consequently the use of water as a biotransformation medium type was dismissed. Despite promising results in shake flasks (Figure 3.14), the biocatalyst performs best in terms of viability and activity when maintained in broth medium.

Media types are pictured in Figure 3.16 post clarification. There is a clear difference between all the types including the 2L and 7L spent broth types, despite the complex broth medium being identical in both cases. Spent broth from the 7L fermenter was distinct from that of the 2L. The fermentation process yields different biocatalyst concentrations (section 3.2.2.2) and this is reflected in the spent media constituents.

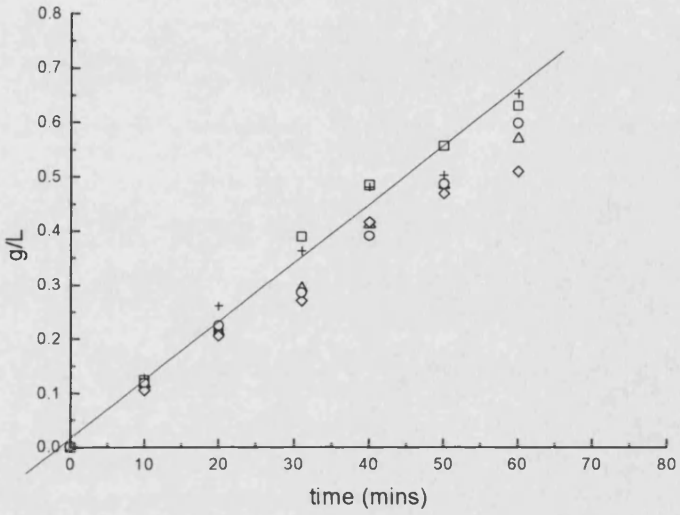


Figure 3.13: Shake flask experiments on various biotransformation media. Lactone produced (g/L) over time. (□) Spent fermentation broth. (◊)Spent fermentation broth (+) Defined. (Δ) Phosphate buffer. (○) Phosphate buffer with biocatalyst washing step.

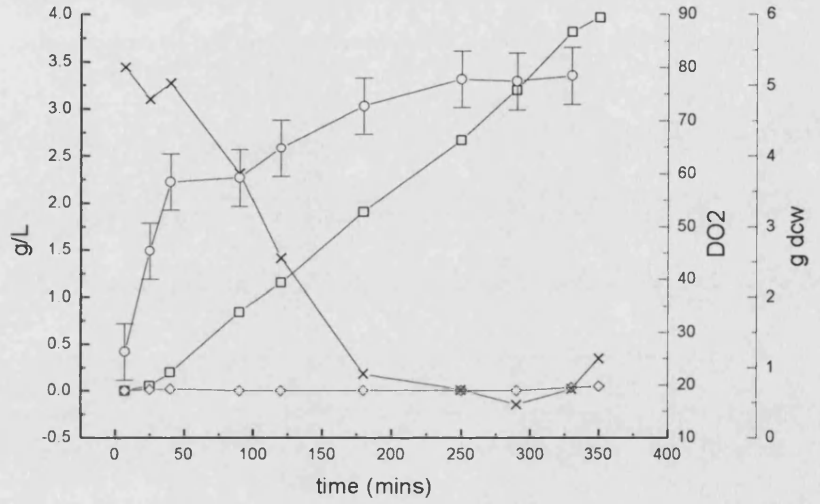


Figure 3.14 Profile of a low (1.5g/L dcw) *E.coli* Top10 [pQR 239] biotransformation carried out in spent fermentation broth. (□) Lactone, (◇) ketone, (x) dO_2, (O) Biomass.

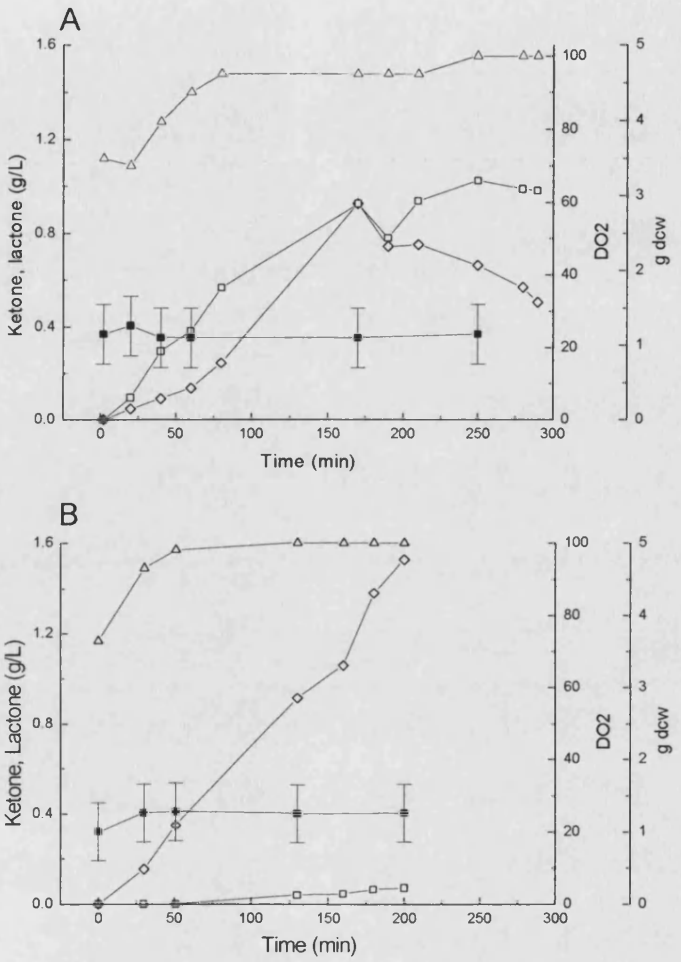


Figure 3.15 Profile of a low *E.coli* Top10 [pQR 239] concentration biotransformation profile carried out in buffer (A) and water (B) medium. (□) Lactone, (◇) ketone, (■) Biomass, (△) DO_2.

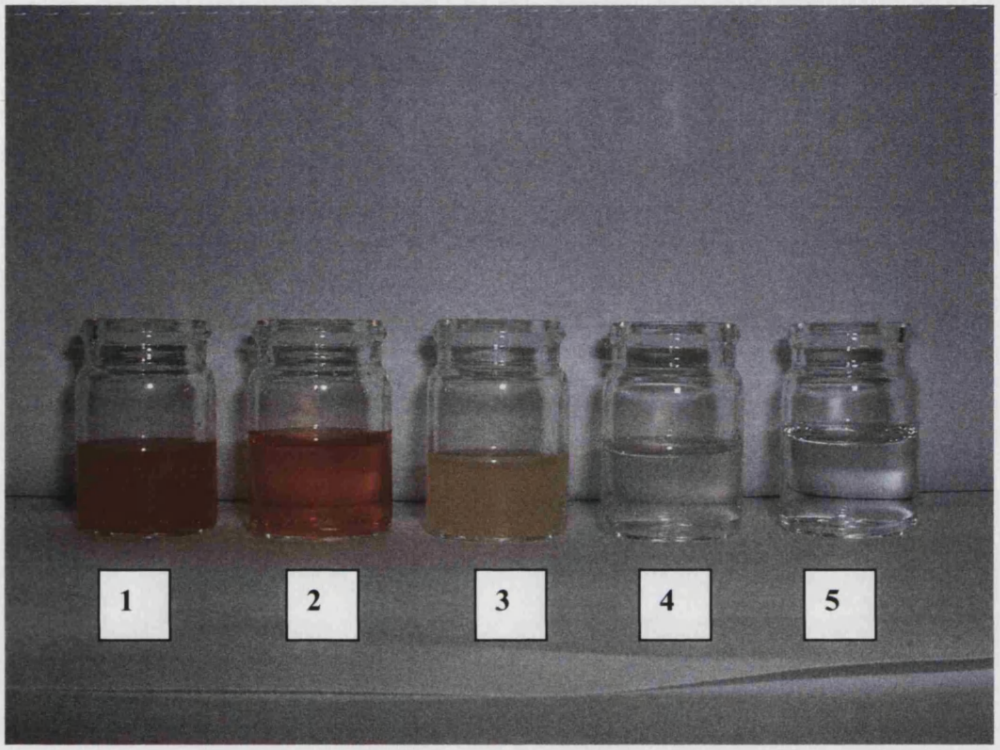


Figure 3.16. Picture of various clarified biotransformation media types: (1) 7L spent fermentation broth; (2) 2L spent fermentation medium type; (3) phosphate buffer medium type; (4) defined medium type; (5) water medium type.

3.3. Discussion

3.3.1. Biocatalyst Production

The production of the CHMO biocatalyst was by way of a 2L or 7L batch fermentation of *E.coli* Top10 [pQR239]. The expression of CHMO during the late log phase of growth achieved 230 to 250 U.g dcw^{-1}, which is approximately half of the 500 U.g dcw^{-1}, expression level previously reported for *E.coli* TOP10 [pQR239] (Doig *et al.*, 2001). The results reported here are the average typical CHMO titre and not optimised expression levels. In addition to the operation of the bioreactor there are differences in the medium used and *E.coli* TOP10 [pQR239] strain variation that could account for the lower CHMO achieved. For example, Doig *et al* (2001) found the percentage of cells not harbouring the plasmid (pQR239) at the end of this fermentation to be approximately 10%; a lack of plasmid stability could be a contributing factor in this case. Further work to optimise cell banking, and the fermentation step is necessary to achieve higher CHMO titres. However, the fermentation step described was acceptable for the purpose of achieving a robust model biocatalyst to undertake the Baeyer-Villiger type oxidation.

The protocol for the 2L (1.5L working volume) fermentation was developed from previously published work (Doig *et al.*, 2001). *E.*coli Top10 (pQR239) fermentation characterisation and scale-up studies (Doig *et al.*, 2003b) were carried out using a similar complex glycerol-based medium. Under the optimum growth and induction conditions, the final biomass yield was 5.5 g.dcw.l^{-1} and the maximum specific activity was 0.9h^{-1}. The protocol developed for this work was based on the work of (Doig *et al.*, 2001) except for the fact that batch growth of *E.*coli Top10 [pQR239] culture was to approximately 4 g.l^{-1} at which point the heterologous expression of CHMO was induced over a 3 hour period. Using this method the final biomass yield was higher at 7.6 and 12 g.l^{-1} for the 2 and 7L fermentations respectively. The 7L (5.3L working volume) process was scaled-up from the 2L process with the key objectives being to achieve a similar final biomass concentration and CHMO titre. To achieve this it was necessary not to limit oxygen transfer

to the culture. The DOT was therefore maintained above 5% during growth and near zero during induction. A near zero DOT during induction was achieved whilst maintaining the maximum OUR level during this period. To achieve the relevant DOT the agitation rate was carefully controlled by a PLC or manually adjusted as necessary.

Initially scale-up of agitation from the 2L to the 7L fermentation was achieved by maintaining tip speed profiles of 1.7 - 2.5 $m.s^{-1}$ (appendix VII) since this method keeps agitator shear-forces constant (Doran, 1998). Similar agitator shear forces between scales were preferable since any change in the constituents of the fermentation broth (caused as a result of shear damage to cells) between batches would result in increased variance during subsequent biotransformation operations. However at constant tip speed, reduced oxygen mass transfer in the larger vessel was restricting culture growth (data not shown) such that growth of cells prior to induction was slower. As a result, scale-up entailed keeping fermentation profiles as similar as possible; this was achieved by following similar stirrer speeds and times as in the 2L vessel. Airflow rates were scaled up based upon volumes of air per volume of broth per minute (vvm) and equalled 0.67 vvm in both examples. All other reagents were scaled-up proportionally based on volume.

The 7L LH2000 fermenter has a larger aspect ratio but a lower ratio of vessel diameter to impeller diameter when compared to LH210_a. Consequently the larger fermenter would be expected to have improved mixing and mass transfer flows, which would contribute to improved growth characteristics. However, higher tip speeds were required at the large scale to achieve a similar specific growth rate and profile. An analysis of fermenter specification is given in Table 3.2, and indicates that the higher tip speeds encountered in LH2000 resulted in improved mixing (higher Reynolds number). It is the improvement in agitation which is responsible for the increased OTR (at the 7L scale), observed indirectly from the OUR profile of the culture growth. However, an undesirable consequence is the increased specific power input which results in a 4 fold increase in the shear rate, as determined by the

calculations given in appendix VIII. This increased shear rate does not seem to have adversely affected the cells.

The amount of extracellular protein observed for the 7L fermenter was more than for the 2L; 0.25 $mg.ml^{-1}$ in the 2L and 0.38 $mg.l^{-1}$ in the 7L fermentation. This difference can be explained by the increased biomass concentrations achieved in the larger fermenter since the mean specific protein concentration is the same for both fermenter cultures at 0.03 $g.gdcw^{-1}$. Regardless of which production scale is used to produce the biocatalyst the cells can be stored in the fridge for a period of 7 days before use.

Parameter	**LH210a**	**LH2000**
Volume (l)	1.6	5.3
Tip speed ($m.s^{-1}$) at N = 1000 rpm	2.1	3.3
Re (minimum)	2393	3556
Pg/V (maximum) (Wl^{-1})	1.1	2.8
Shear rate (s^{-1})	476	1963

Table 3.2. Fermenter parameters were calculated from equations detailed in appendix VII and appendix VIII. It was necessary to assume broth density (1100 $kg.m^3$), and the viscosity (0.01 Pa.s) of the fermentation medium.

3.3.2. The Model Biotransformation

Though the whole-cell bioconversion is possible in a number of media types a rapid loss of biocatalyst stability was observed when using phosphate and water medium. A loss of viability limits the period during which maximum activity can be maintained during a biotransformation reaction. Maintaining maximum volumetric activity is required if the yield of combined lactone is to be optimised. The whole-cell Baeyer-Villiger type oxidation was

representative of other biotransformations in that the biocatalyst viability is diminished under operating conditions (Chen *et al.*, 2002). Maximum biocatalyst activity is limited to a period of 4 hours (Figure 3.7) under typical biotransformation conditions. It might be the case that an improvement of biocatalyst stability can be affected by a change in the operating conditions; for instance temperature. Nevertheless, there will be a finite period during which the process can be operated efficiently.

The model biotransformation is also constrained by both substrate and product inhibition. The effect of inhibition was demonstrated (Figure 3.11) during a bioconversion in a STR. The production of lactone peaked at 175 minutes as ketone concentration reached inhibitory levels. This example used a strategy of continual feeding of the substrate. To develop the process a strategy of controlling the ketone feed so as not to over or under feed substrate to the reactor was proposed. This would eliminate the problem of substrate inhibition, and improve the yield of lactone produced per gram of substrate used. The model Baeyer Villiger biotransformation like many potentially important biotransformations cannot be implemented as economically viable processes because of low volumetric productivities (Cheetham, 1994). In many cases this constraint on the bioconversion is due to accumulation of product and or substrate during the bioconversion (Chauhan *et al.*, 1997). In most cases, the biocatalyst may be inhibited by the reaction substrate and products at the biochemical level through a natural feedback mechanism aimed at controlling metabolic activity (Liese *et al*, 1993). Substrate and product inhibition for the Baeyer-Villiger monooxygenase biotransformation model was previously described. The optimum ketone concentration was between 0.2 and 0.4 gl^{-1} and at product concentrations above 4.5 to 5gl^{-1} the specific activity of the whole cells fell to zero (Alphand *et al.*, 2003).

The variation in yield of biocatalyst and therefore variation in constituents of the spent medium between fermentation scales will be transferred to any subsequent biotransformation step. Such variation is not desirable for monitoring by NIR spectroscopy. To limit these effects the cells can be

resuspended in fresh biotransformation medium. Biotransformation medium exchange was expected to enable the concentration of biocatalyst, and glycerol to be controlled whilst reducing the variability of fermentation products including protein, dissolved in the medium. The results, however, indicated lower viability and activity of the biocatalyst in alternative media types, whilst at low concentration in complex medium *E.coli* Top10 [pQR239] was observed to grow. To minimise variation in biotransformation sample sets for NIR spectroscopic analysis the complex spent fermentation medium type was used, from either the 2 or 7L scales.

3.4. Conclusion

- *E.coli* Top10 [pQR239] biocatalyst production was undertaken at a 2L and 7L scales. The developed procedure included the batch growth of *E.*coli Top10 [pQR239] culture to approximately 5g.l^{-1} at which point the heterologous expression of CHMO was induced over a 3 hour period.

- Though both the 2L and 7L fermentation methods produced similar specific intrinsic CHMO concentrations, the total yield of cells and therefore yield of CHMO differed between scales.

- The larger 7L scale fermenter required increased agitator tip speeds to achieve a similar DOT and cell growth profile. The increased tip speed meant greater power input, which contributed to a marked increase in the concentration of *E.coli* Top10 [pQR239].

- Stored biocatalyst suspended in spent fermentation medium in a refrigerator at 4°C is stable for 170 hours without a reduction in the CHMO titre. The biocatalyst was shown to be stable under operating conditions for 4 hours after which a loss in CHMO activity was observed.

- All routine biotransformations require glycerol as a reductant.

- Baeyer-Villiger bioconversions of ketone to combined lactone were achieved using the method described in section 3.1.2. To maximise the volumetric activity of the biocatalyst, control of substrate feeding is required.

- The model biotransformation was successfully carried out on a number of biotransformation media types. The complex medium type is considered the least favourable for NIR spectroscopic analysis but the most favourable for the biocatalyst in-terms of activity and viability.

4. Near Infrared Spectroscopic Monitoring of the Biotransformation

This chapter examined the application of NIR spectroscopy to monitoring the whole-cell Baeyer-Villiger biotransformation of a cyclic ketone to regio-isomeric lactone products. The NIR instrument was employed at-line (rapid off-line) and on-line (*in situ*). At-line analysis required samples to be withdrawn from the bioreactor before the sample contents (matrix) could be analysed. On-line analysis interfaced the analyser directly to the bioreactor contents (matrix) and as such, the bioreactor did not require samples to be withdrawn. The development of a NIR sampling method and subsequent calibration of spectra to the GC reference method was pursued through a feasibility study that examined spectral responses to known samples.

4.1. Materials and Methods

4.1.1. NIR Spectral Sampling

A fibre optic transmission immersion probe (FOSS UK Ltd, Warrington, Cheshire, UK), with a total path length of 1.5 mm was used for all at-line and on-line sample analysis. During on-line work, the probe was fitted to the head-plate of the reactor through a 20 mm port and was situated so that the cell-path was 3.0 cm from the reactor base. A Foss NIR systems model 6500 visible/NIR spectrophotometer equipped with a halographic concave diffraction grating, and a lead sulfide (PbS) detector was used to measure the spectra (Figure 4.0). VISION software (FOSS UK) was used with a PC workstation to control spectra collection. On-line samples were ratioed to air background at regular intervals using 64 co-averaged scans per sample throughout the bioconversion process. At-line sampling required a 5 ml volume of broth to be removed from the reactor at regular intervals and immediately centrifuged at 10,000 rpm for 2 minutes (section 3.1.2.2) before NIR spectra (referenced to air) were recorded unless stated otherwise. Samples taken at-line consisted of 32 co-averaged scans, unless stated

otherwise. Off-line spectra of lyophilised lactone[1] were ratioed to a ceramic background using a Foss NIR systems reflectance probe.

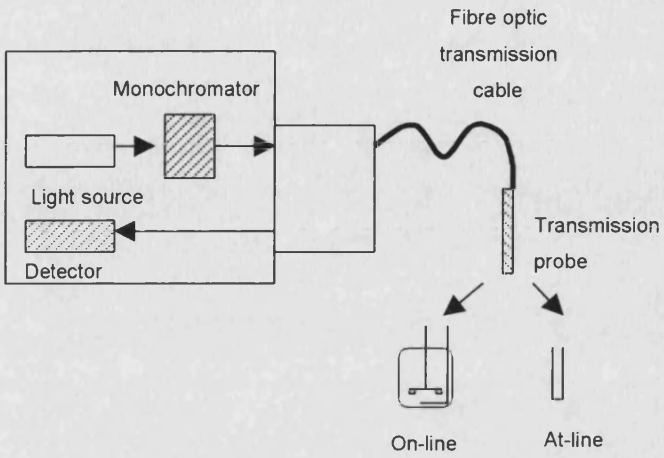


Figure 4.0 The NIR system for monitoring the bioconversion.

4.1.2. NIR Multivariate Calibration

NIR spectroscopy is an optical technique that involves measurement between the visible (700 nm) and the mid-IR (2200 nm) region. In quantitative NIR spectroscopy, empirical relationships (calibration models) are derived between the absorption (or transmission) of NIR radiation and the corresponding reference assays for the variable of interest. Calibrations between 2nd derivative (segment size of 20 nm and gap size of 0 nm) mathematically treated spectra and associated reference analytical data were produced using a partial least-square (PLS) approach. Calibration was undertaken in the 1550 to 1820 nm region of the electromagnetic spectrum. The PLS method employs factor analysis and then uses a subset of the resulting factors to complete the regression equation. The best possible regression models were achieved by a cross validation procedure optimising the number of factors such that either a minimum prediction residual error sum of squares (PRESS) is reached or the fewest factors required to achieve a standard error of calibration (SEC) similar to the associated gas

chromatography standard error (Appendix IX). Cross-validation involved each calibration sample being dropped out of the calibration set, generating regression equations on the reduced calibration set and then subsequently predicting the value of the dropped standard (Chalmers *et al.*, 1997). VISION 2.22 software (FOSS UK Ltd, Warrington, Cheshire, UK) was used in conjunction with a PC workstation for the storage and reduction of all data. Using the regression models, quantitative predictions of variables were made on sets of unknown process samples.

4.1.3. Qualitative NIR Analysis

For qualitative analysis, data was displayed using VISION 2.22 software. The data matrix for any given spectra or 2^{nd} derivative mathematically treated spectra was able to be exported from VISION 2.22 where necessary.

Spectral sample data from VISION 2.22 was exported as a text file to MatLab 5.3 (Math Works, MA, USA) where the data matrix was transposed before exporting as an ASCII file.

4.2. Results

4.2.1. Feasibility Study

The feasibility study was designed to provide information on the absorbances expected in the process sample matrix and to demonstrate that a quantitative analysis is possible. The process sample matrix of undefined biotransfomation medium type is highly complex and so it is not possible to scan all the ingredients in the mixture independently. However, by knowing the unique absorbance of the substrate and product material a focus on where in wavelength terms to perform the NIR analysis is provided. Spiking samples with analytes was used to provide a range of absorbances, which were used to check for NIR response linearity and accuracy. Finally, the feasibility study was used to decide on the regression method to be used and on how to present and take samples using the NIR system.

4.2.1.1. NIR Spectra of Substrate and Products

NIR absorption bands are associated with the overtone's and combination bands of the –CH, -NH, and –OH fundamental vibrational absorptions that occur in the mid-infrared region. The overlapping absorptions found in the NIR region are weaker in intensity than comparable mid-IR bands (Brimmer *et al.*, 1996, Stallard, 1997). Thus, chemometric techniques are required in quantitative analysis of NIR spectra (Vaidyanathan *et al.*, 1999b). The CHMO biotransformation was undertaken in aqueous fermentation broth, resulting in some relevant information being masked by strong water absorption bands as can be seen in Figure 4.1.

The substrate (ketone) was analysed using a NIR spectrophotometer fitted with a transmission probe. An enantiomer of the lyophilised lactone[1] was also analysed using NIR spectroscopy fitted with a reflectance probe. The ketone and lactone[1] spectra were compared to that of water as shown in Figure 4.1. Lactone[1] is commercially available but the combined lactones ((-) 1(S), 5(R)2-oxabicyclo(3.3.0)oct-6-en-3one and (-) 1(R), 5(S) 3-oxabicyclo(3.3.0)oct-6-en-2-one) are not. Lactone[1] was therefore used during

feasibility analysis to locate wavelength regions for calibration of the ketone and lactone. As the regio-isomers of combined lactone are distinct molecules with different physical properties, individual IR vibrational signatures would be expected. However, because the regio-isomeric disparity is only the site of oxygenation on the cyclic-ketone (Shipston *et al.*, 1992) no distinction of these forms was made during NIR analysis. The NIR region contains first order, second order and combinations of fundamental vibrations occurring in the mid-IR region, resulting in broad overlapped spectral bands such that the spectra of lactone[1] was able to provide a good indication of the wavelength region for combined lactone calibration.

The spectra of both the ketone and lactone[1] (Figure 4.1) are similar with significant peaks in the first overtone region 1666nm, 1750nm, 1900nm and 1950 nm. The combination band region also contains peaks including 2150 nm and the region 2250-2300 nm. The majority of the vibrations result from CH - CH_2 bond stretches that are common to both the ketone and the lactone. It was therefore apparent that selection of unique spectral features for calibration was not possible. It is because of the chemical complexity of the biotransformation processes and the lack of distinct spectral features of analytes that it is difficult to relate spectral variation observed in the broth spectra to changes in concentration levels of analytes (Arnold *et al* 2002). A multivariate partial least squares (PLS) method was therefore used during quantitative calibration development (Osborne *et al.*, 2002). Distinctive peaks of water at 1450 nm and 1950 nm limit the wavelength region available for calibration. This is because the wavelength regions at which water absorbs result in overwhelming water absorption bands such that no or very little electromagnetic radiation in this region reaches the detector. In addition, the use of fibre-optics to facilitate on-line monitoring limits the calibration region further as the fibre-optic cable absorbs NIR radiation above 2200 nm. This is seen in Figure 4.1 as the absorption peak at 2400nm.

Since absorption of NIR rays by biotransformation medium is dominated by a water band, only the region of 1500 to 1850 nm, between the peaks was applied to the monitoring of ketone and combined lactone. To distinguish

further the absorption spectra of ketone and lactone[1] the spectra in Figure 4.1 was converted to the 2nd derivative (Figure 4.2). A 2nd derivative absorption spectra were used instead of raw spectra to reduce the effects of baseline differences (Brimmer *et al.*, 1993) and to improve the resolution of distinctive wavelength features (Arnold *et al.* 2002). Figure 4.2 indicates differences between the ketone and lactone analytes absorption bands in the wavelength region 1650 to 1750 nm. It isn't possible to precisely identify bands associated with the analytes of interest using Figure 4.2, as the absorption bands of the analytes in aqueous solution will differ from the corrseponding raw material (Osborne *et al.*, 2002). However, the results enable a wavelength region to be assigned on which to develop quantitative calibration models.

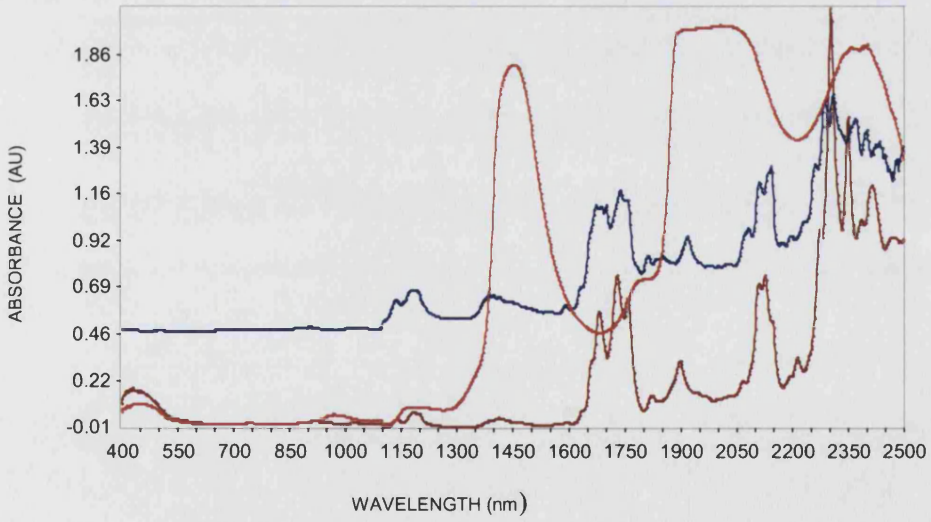


Figure 4.1 NIR reflectance spectra of lyophilised lactone[1] (blue) referenced to a ceramic cell. NIR transmission spectra of ketone (brown) and water (red) referenced to air.

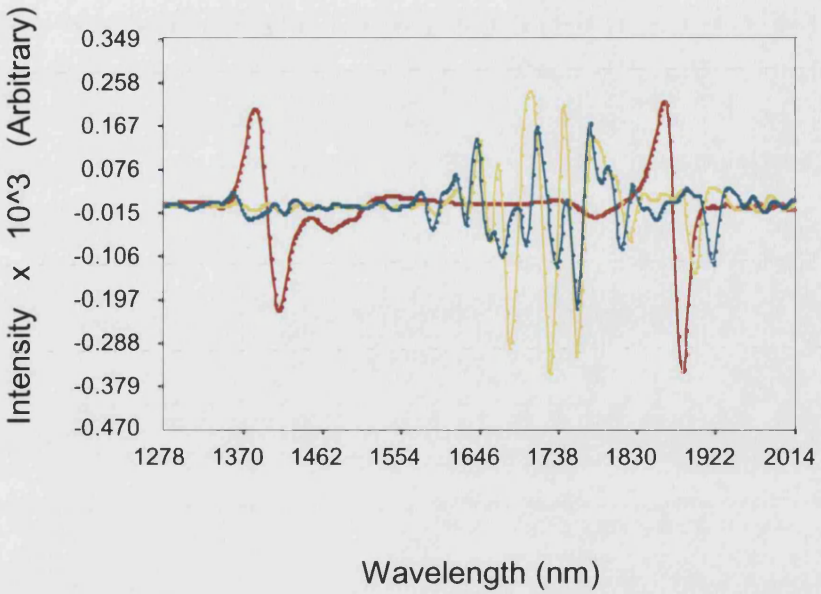
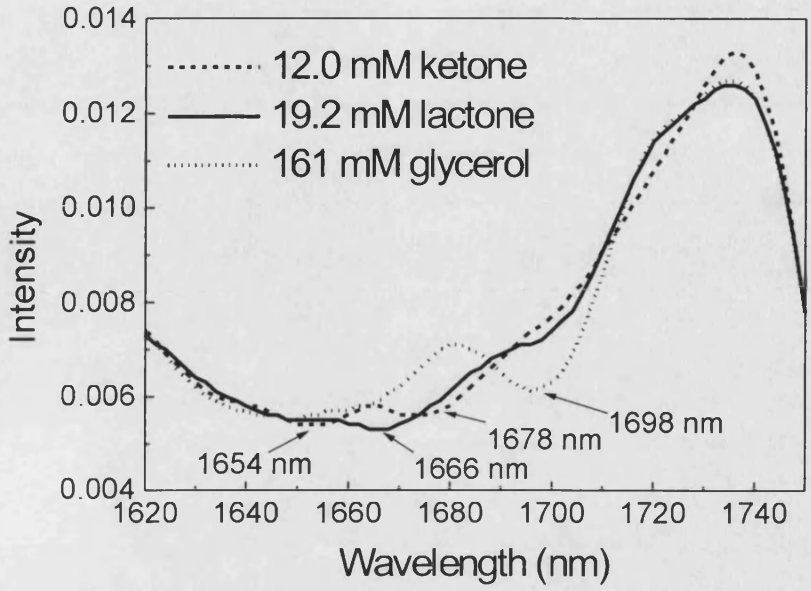


Figure 4.2 A 2nd derivative NIR reflectance spectra of lyophilised lactone[1] (blue) referenced to a ceramic cell. NIR transmission spectra of ketone (yellow) and water (red) referenced to air.

4.2.1.2. NIR Spectra Background

In order to identify the characteristic absorptions of the individual components in the biotransformation medium, the NIR spectra of a series of ketone, lactone and glycerol concentrations were collected. Glycerol was included since this was a major constituent of all biotransformation media types. Figure 4.3 shows the second-derivative spectra of ketone at a concentration of 1.3 and 13 gl^{-1} (12.0 & 120mM), lactone at 2.4 and 12 gl^{-1} (19.2 & 96.0 mM), and glycerol at 15 and 30 gl^{-1} (161.3 & 322mM), respectively. The key features of the spectral region; 1600-1740nm are attributed to the C-H first overtone stretching vibration. The absorption at around 1654, 1678, and 1720 nm are mainly due to ketone. Two absorption regions were observed at around 1666 and 1702 nm, which depend on lactone. The absorption of glycerol was observed at 1700 nm.



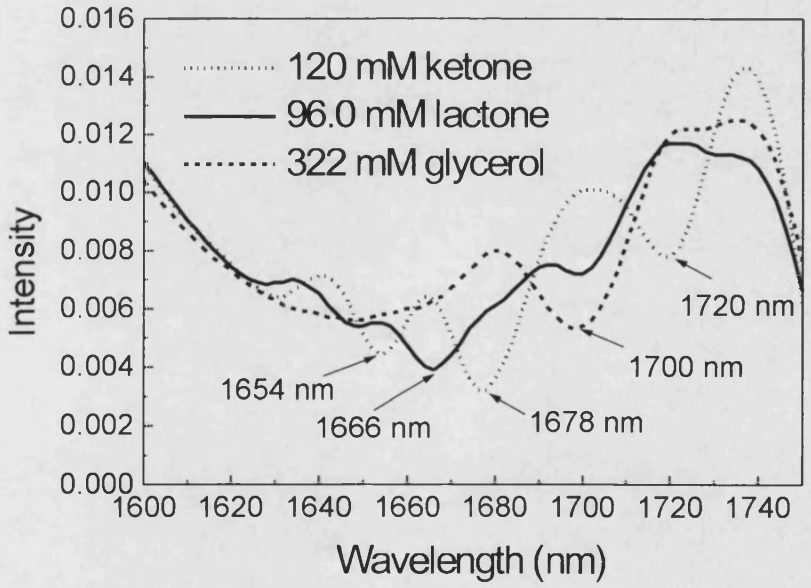


Fig. 4.3 The second derivative NIR spectra (referenced to water) of ketone, lactone and glycerol samples.

4.2.1.3. Amplitude Response

Identification of absorption bands of interest associated with the analytes of interest was achieved by spiking lactone and ketone into the biotransformation matrix (section 4.2.1.2). This procedure was then used to investigate the linearity of the NIR response taking a wavelength value of 1664 nm and 1678 nm for lactone (Figure 4.4) and ketone (figure 4.5) respectively. A plot of analyte concentration at the wavelengths of interest (Figure 4.6) describes a linear response in both cases with a standard deviation of 0.14 gl^{-1} for lactone and 0.13 gl^{-1} for ketone.

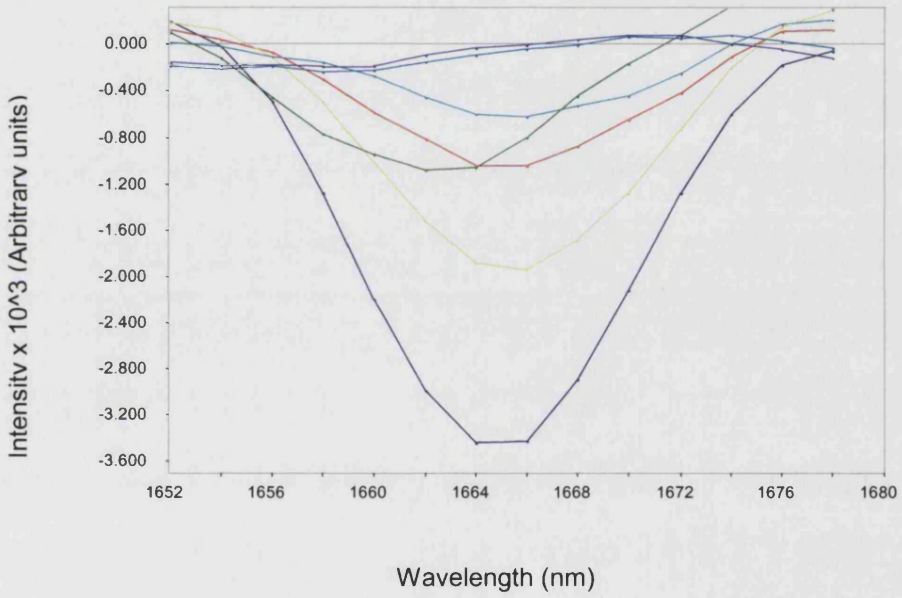


Figure 4.4: A 2nd derivative NIR transmission scans of aqueous lactone[1] samples referenced with RO H_2O. Including: green = 0.625 gl^{-1}; pale blue line = 1.25 gl^{-1}; red line = 2.5 gl^{-1}; yellow line = 5 gl^{-1}; blue line = 10 gl^{-1}.

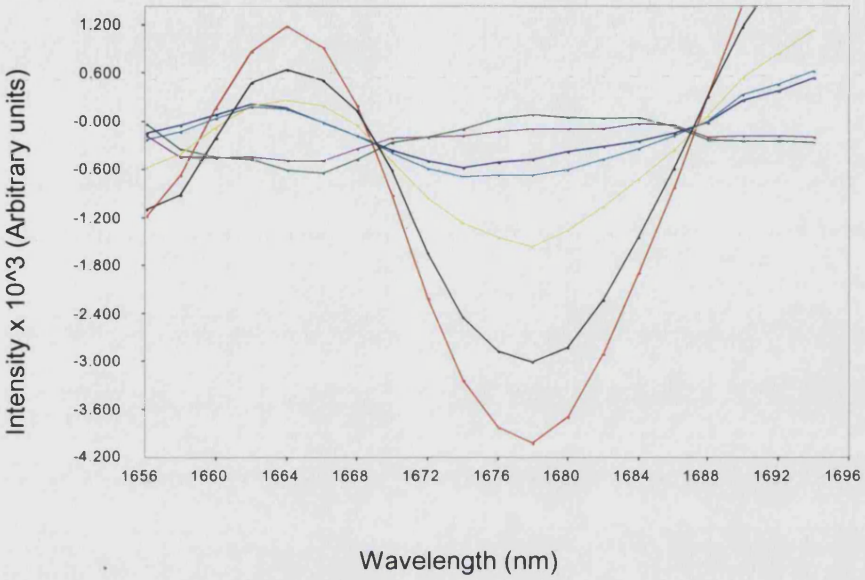
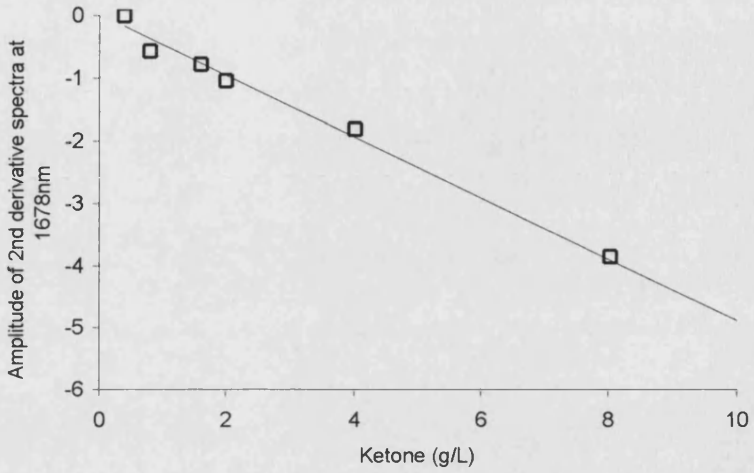


Figure 4.5: A 2nd derivative NIR transmission scans of aqueous ketone samples referenced with RO H_20. Including green = 0.4 gl^{-1}; pink = 0.8 gl^{-1}; dark blue = 1.6 gl^{-1}; light blue = 2.0 gl^{-1}; yellow = 4.0 gl^{-1}. black = 8.0 gl^{-1}; red = 10 gl^{-1}.

(A) R= 0.99. SD=0.13 units



(B) R= 0.99. SD=0.14 units

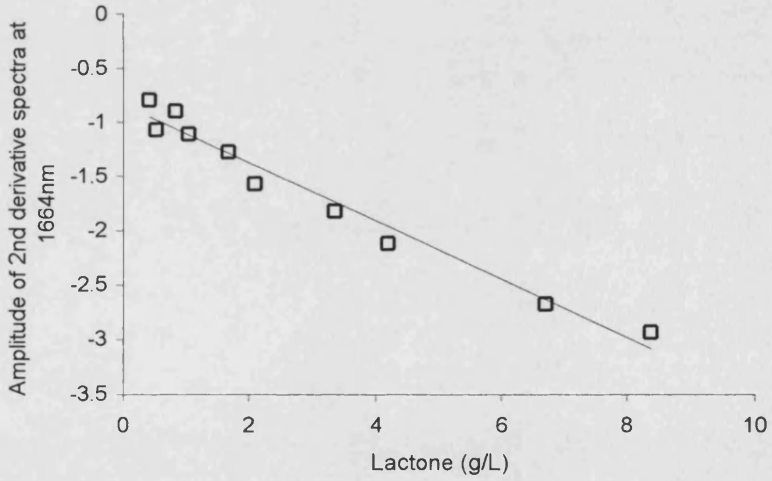


Figure 4.6 (A): amplitude linearity of a 2nd derivative NIR signal at 1678 nm for ketone. (B); amplitude linearity of 2nd derivative NIR spectra at 1664 nm for lactone.

4.2.1.4. NIR Response Time

Spectra were collected using a fibre-optic transmission probe (section 4.1) rather than a fixed cell. The advantage of using a transmission probe is that it can be positioned to sample either at-line or on-line. For each method, a sufficient number of sample absorption scans was co-averaged to obtain a suitable signal to noise level (accuracy) for quantitative monitoring. However, as the number of co-averaged scans increase so does the sampling time (Figure 4.7). At 1.5 minutes the sample time to co-average 128 scans is fast when compared to the 35 minutes to prepare and run a GC sample (section 3.1.3.5). However, at-line NIR sampling involves removing a volume of process fluid from the reactor and clarifying it by centrifugation (section 4.1). The cell-free at-line sample is then presented to the NIR transmission probe. The time taken to sample and clarify an at-line sample took 4 minutes, in addition to the NIR spectral acquisition time. On-line sampling was achieved with the transmission probe in the reactor, so sampling was equal to the time taken to collect and co-average the required number of scans.

NIR monitoring of the biotransformation system was used to enable an improvement in biotransformation process control. To achieve this, the monitoring system needed to be both reliable and rapid. It was therefore necessary to collect sample spectra using an optimal method of sample presentation. For on-line sampling, increasing the number of co-averaged scans significantly reduced the standard deviation (SD) associated with the signal reproducibility (Figure 4.8). The optimum signal to noise ratio for on-line sampling was calculated as being 64 co-averaged scans, giving a sample time of approximately 50 seconds. The at-line sample signal reproducibility was not significantly improved (Figure 4.8) and so to limit the total at-line sample time to <5 minutes, 34 scans were co-averaged. Nevertheless the signal reproducibility of on-line NIR analysis was not as accurate as that of at-line (Figure 4.9). A loss of signal reproducibility is evident when the on-line water spectra are compared to the at-line results (Figure 4.9b).

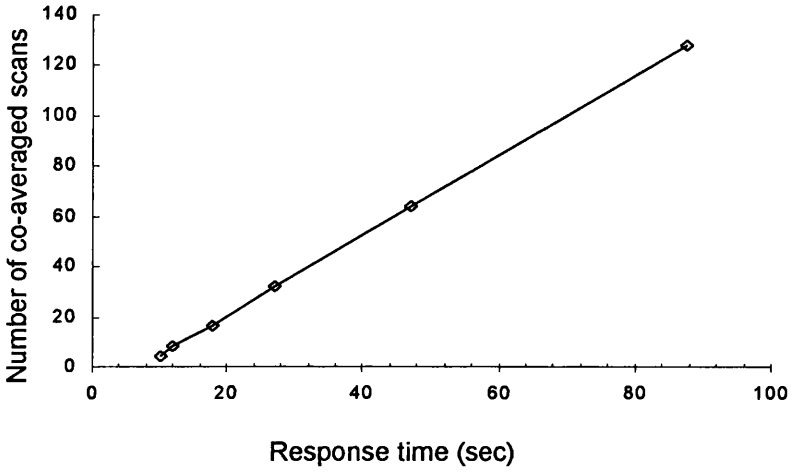


Figure 4.7 NIR transmission scan time vs number of co-averaged scans.

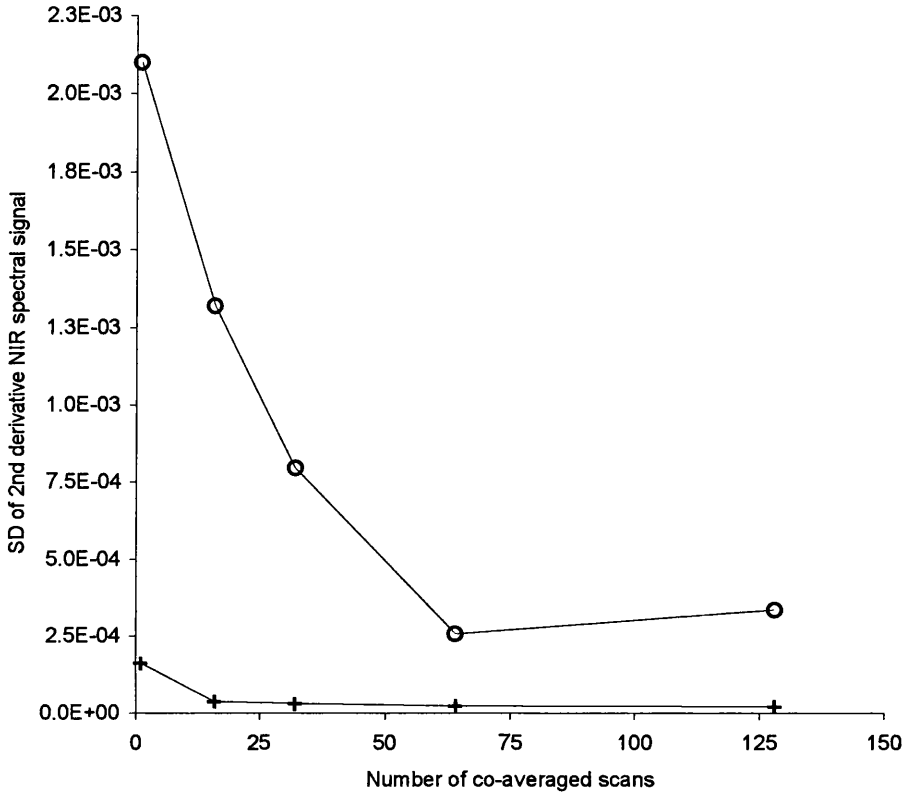


Figure 4.8 Increasing NIR scan number improves signal reproducibility. The mean SD was calculated between transmission spectra of water between 1550 to 1820 nm. (0) on-line in a STR at 37 °C with agitation, stirrer speed 800 rpm. (+) at-line sampling.

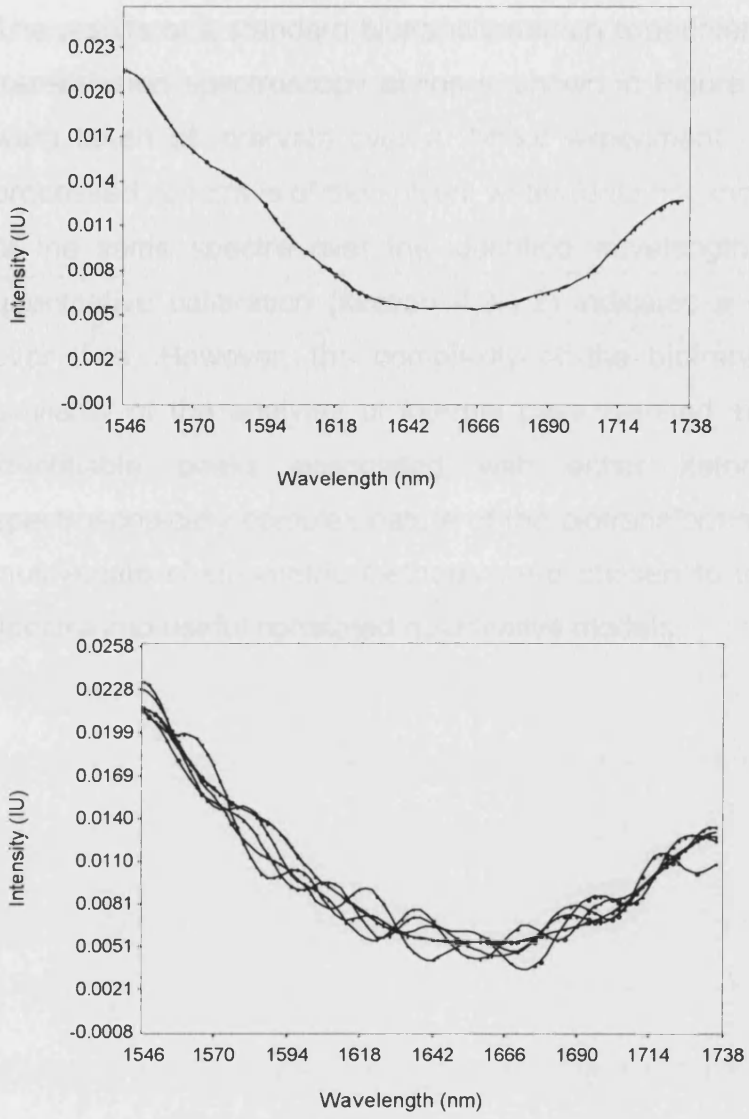


Figure 4.9. (A): Four 2nd derivative transmission spectra of H_20 sampled at-line. (B): Four 2nd derivative transmission spectra of H_20 sampled at-line and four 2nd derivative transmission spectra of H_20 sampled on-line. Intensity of the NIR spectra has arbitrary units (IU). On-line sampling was in a STR with aeration at 0.67 vvm.

4.2.1.5. Profile of Fed-batch Baeyer-Villiger Bioconversion

The results of a standard biotransformation experiment monitored using NIR transmission spectroscopy at-line is shown in Figure 4.10. The 20 samples were taken at intervals over a 3-hour experiment. The shape of the un-processed spectra is of the solvent water (data not shown). The 2nd derivative of the same spectra over the identified wavelength region of interest for quantitative calibration (section 4.2.1.2) indicates a change in the spectra over time. However, the complexity of the biotransformation matrix and similarity of the analytes of interest have resulted in broad bands with no identifiable peaks associated with either ketone or lactone. The spectroscopically complex nature of the biotransformation matrix means that multivariate chemometric methods were chosen to translate these primary spectra into useful correlated quantitative models.

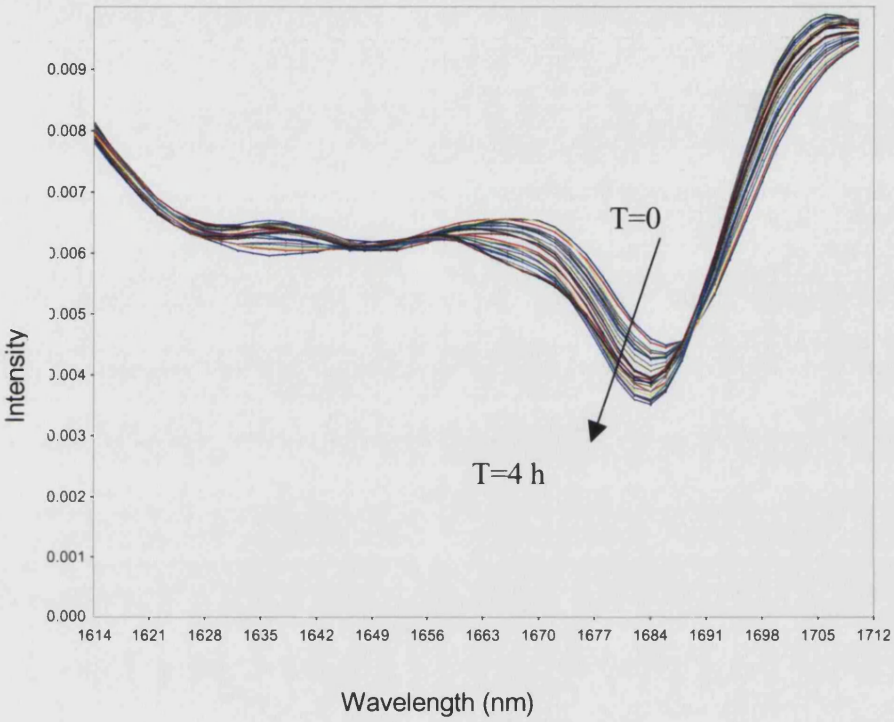


Figure 4.10. At-line 2nd derivative NIR spectra profile of a series of 20 biotransformation samples taken at-line and referenced to air – calibration wavelength region. T = time (h).

4.2.2. Quantitative Analysis

The NIR spectral region is situated between the end of the visible spectral region at 700nm and the beginning of the fundamental IR spectral region at 2500nm. Quantitative NIR spectroscopic empirical relationships (calibration models) were derived between the absorption (or transmission) of NIR radiation and the corresponding reference assays for the variable of interest. The calibration was then used to predict the properties of unknown process samples from the output of the NIR spectrophotometer. Both the construction and prediction stages are vital in the generation of calibration models for NIR (Osborne *et al.*, 2002). This section reports the first example of the application of fibre-optic NIR spectroscopy for the monitoring of analytes at low concentrations during a whole-cell *E.coli* bioconversion (Bird *et al.*, 2002a). Calibration for substrate and product concentrations both at-line and on-line is discussed.

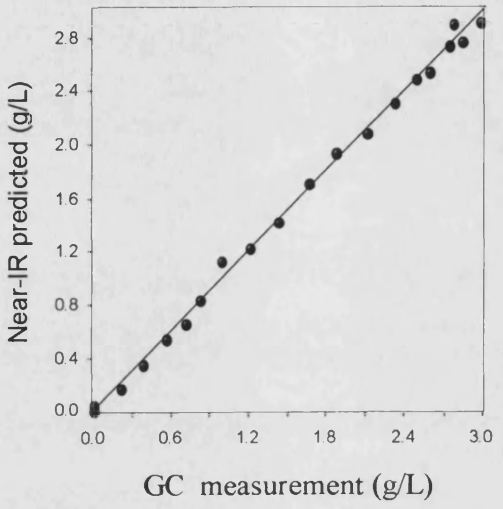
4.2.2.1. Quantitative PLS Model Building

Production of the whole-cell *E.coli* biocatalyst (Doig *et al.*, 2001, Doig *et al.*, 2002) was by way of fermentation (section 3.1.1) prior to utilisation for bioconversion (section 3.1.2). In this example the Baeyer-Villiger reaction was undertaken in the fermentation broth. Any batch variation from the fermentation step was transferred to the subsequent bioconversion and resulted in increased batch to batch variation in the bioconversion step (section 3.2.1). Variation between three separate bioconversions was included in the calibrations, which incorporated 52-53 spectra taken both at-line and on-line. Calibration models were built for ketone and lactone using a PLS method (section 4.1.2). Predicted Residual Error Sum of Squares (PRESS) analysis was used to optimise the equations for number of principle component factors (Appendix IX).

4.2.2.2. Calibration Wavelength Selection

Wavelength regions to be used during calibration of the process sample sets were selected by generating calibration models on a set of 23 samples spiked with ketone and lactone[1]. The wavelength region was expanded iteratively to improve the SEC. The region with the optimal SEC was chosen as the wavelength region of choice for PLS analysis on process samples. Figure 4.11 shows two results of this method during analysis of ketone and demonstrates how the SEC for ketone was improved by careful selection of the wavelength region.

A


Near-IR predicted (g/L)
0.0
0.4
0.8
1.2
1.6
2.0
2.4
2.8
0.0
0.6
1.2
1.8
2.4
3.0
GC measurement (g/L)

B

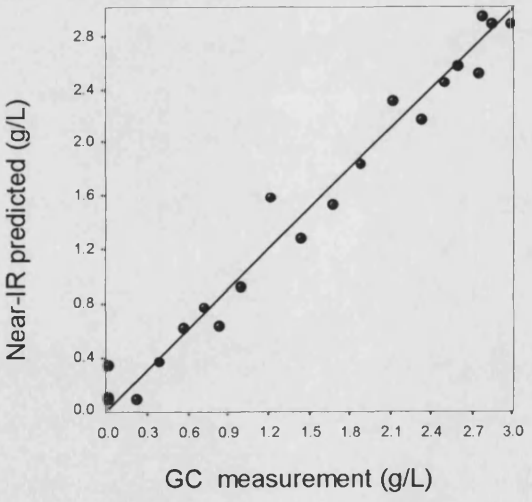

Near-IR predicted (g/L)
0.0
0.4
0.8
1.2
1.6
2.0
2.4
2.8
0.0
0.3
0.6
0.9
1.2
1.5
1.8
2.1
2.4
2.7
3.0
GC measurement (g/L)

Figure 4.11. Parity plots of calibration model for ketone over 2 wavelengths (A) 1550 – 1820 nm, (B) 1500 – 2100 nm. Models were constructed using 2nd derivative spectra of spiked water samples referenced to water, n = 23. Models are based on 3 factors. Model (A) SEC = 0.056 gl^{-1}, (B) SEC = 0.172 gl^{-1}.

4.2.2.3. NIR Calibration of the Model Biotransformation

Calibration of predicted and measured (gas chromatography) concentrations for lactone and ketone using data taken on-line from three bioconversions resulted in standard errors of calibration (SEC); equation A, 0.111 gl^{-1} (1.02 mM) ketone (Figure 4.12) and equation B, 0.182 gl^{-1} (1.452 mM) lactone (Figure 4.13). Validating the on-line equations with a fourth, independent sample data set revealed SEP of 0.130 gl^{-1} (1.200 mM) ketone (Figure 4.12) and 0.177 gl^{-1} (1.425 mM) (Figure 4.13) combined lactone. The SEC of at-line spectra to gas chromatography data was significantly lower than on-line values with equation C, 0.074 gl^{-1} (0.689 mM) ketone (Figure 4.14) and equation D, 0.119 gl^{-1} (0.958 mM) lactone (Figure 4.15). Validation of at-line equations for ketone and lactone resulted in SEP of 0.088 gl^{-1} for ketone and 0.101 gl^{-1} for combined lactone. Figure 4.16 and Figure 4.17 compare the predicted and measured concentration profiles for ketone and lactone. The calibration methods are outlined in Table 4.0 and the validation results in Table 4.1.

Equation	Analyte	n	wavelength (nm)	Factors	R^2	SEC (gl^{-1})	F
A	Ketone	53	1510-1800	4	0.955	0.111	254
B	Lactone	52	1550-1820	4	0.952	0.182	240
C	Ketone	52	1510-1800	5	0.983	0.074	522
D	Lactone	52	1550-1820	3	0.979	0.119	728

Table 4.0: Calibration of NIR spectra to gas chromatography data. All spectral data was mathematically treated; 2^{nd} derivative before PLS regression to build online equations A & B, and atline equations C & D. Three independent Baeyer-Villiger biotransformation data sets were used to build equations. A fourth set was used to perform an external independent validation (Table 4.1).

Equation	Sampling	n	SEP (gl^{-1})	R^2
A	On-line	20	0.1296	0.907
B	On-line	21	0.1767	0.967
C	At-line	16	0.0875	0.957
D	At-line	19	0.1095	0.991

Table 4.1: Validation of NIR calibration models was performed using an external validation set that was independent of the calibration set.

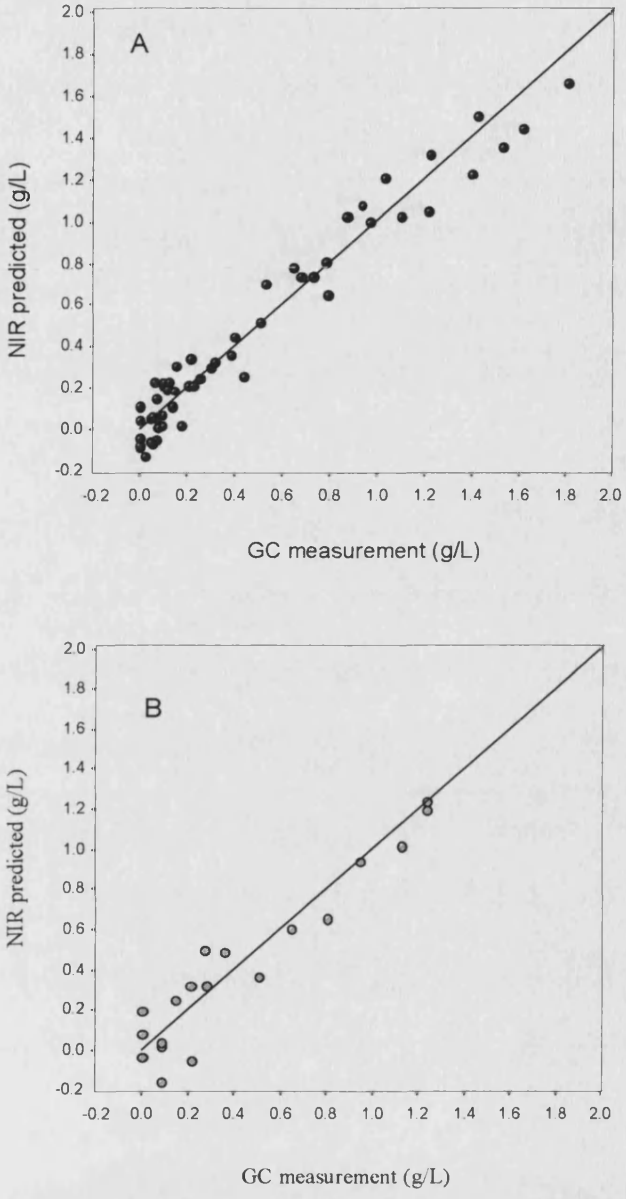


Figure 4.12. Concentration correlation plots. (A) Online ketone calibration equation **A**. SEC = 0.111 gl^{-1}. (B) Prediction of unseen online data using equation A. SEP = 0.130 gl^{-1} external validation method.

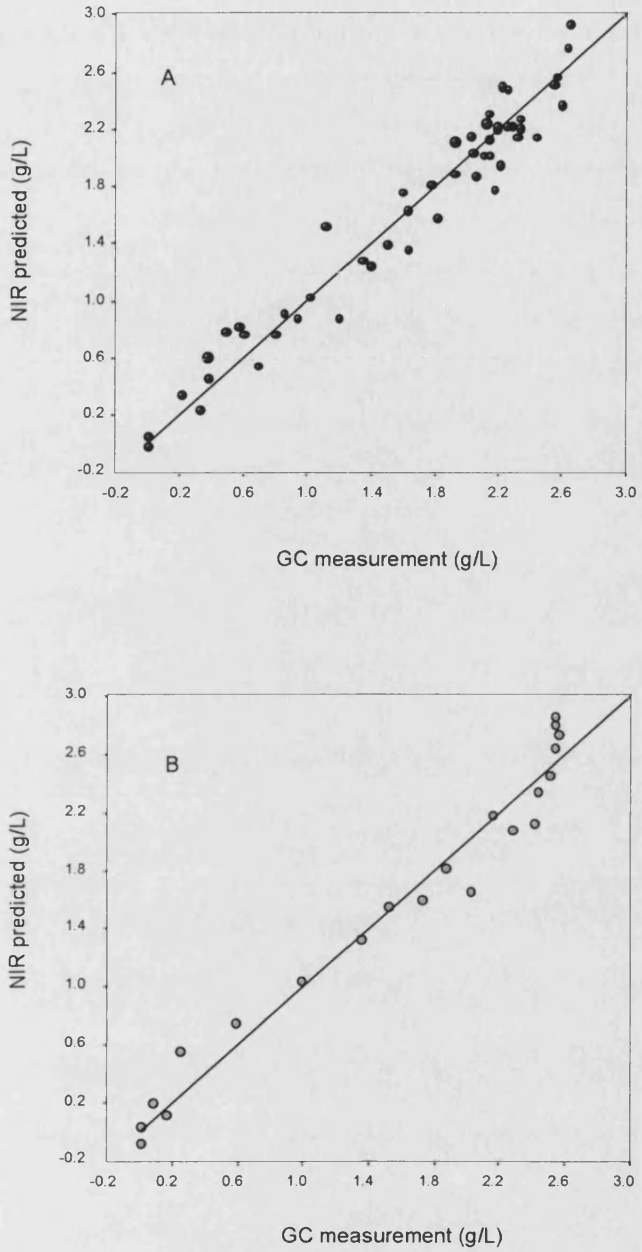
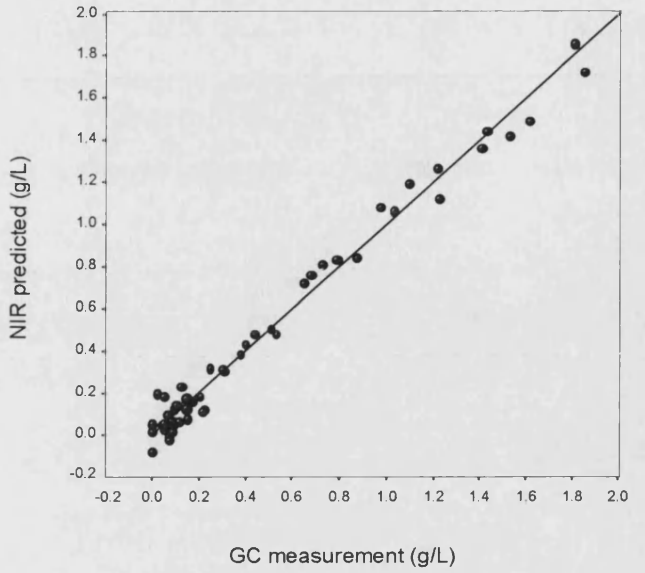


Figure 4.13. Concentration correlation plots. (A) Online combined lactone calibration equation **B**. SEC = 0.182 gl^{-1} (B) Prediction of unseen online data using equation **B**. SEP = 0.177 gl^{-1} external validation method.



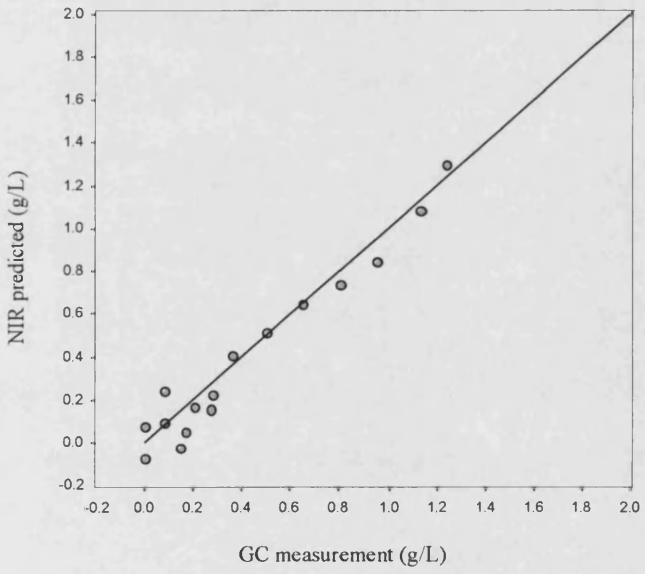


Figure 4.14 Concentration correlation plots. (A). At-line ketone calibration equation **C**. SEC = 0.074 gl^{-1}. (B). Prediction of unseen at-line data using equation **C**. SEP = 0.088 gl^{-1}, external validation method.

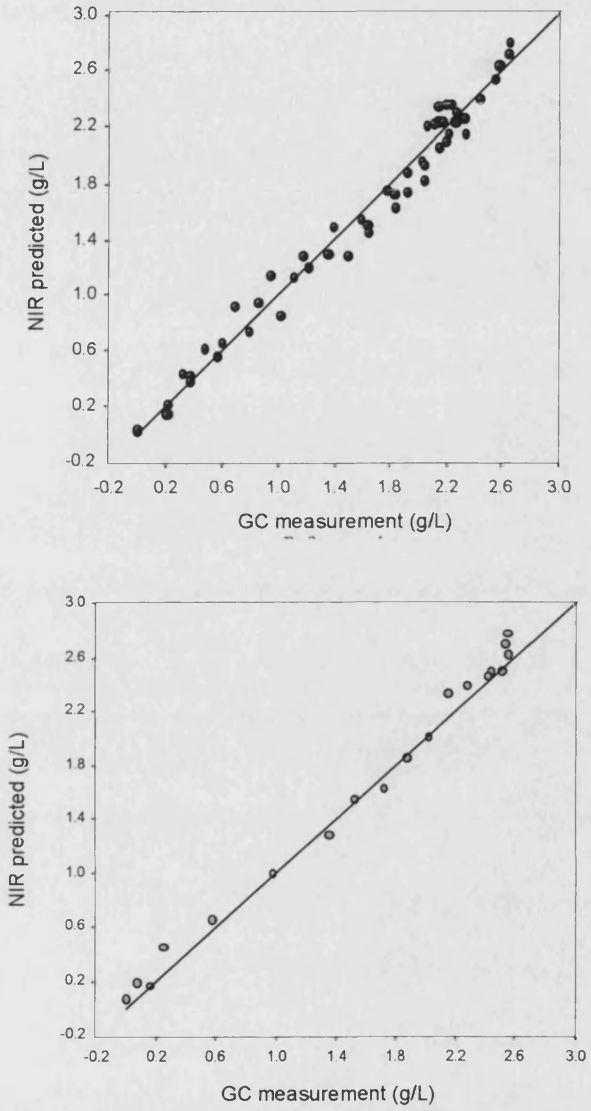


Figure 4.15. Concentration correlation plots.(A) At-line combined lactone calibration equation **D**. SEC = 0.119 gl^{-1}. Prediction of unseen at-line data using equation **D**. SEP = 0.110 gl^{-1}, external validation method.

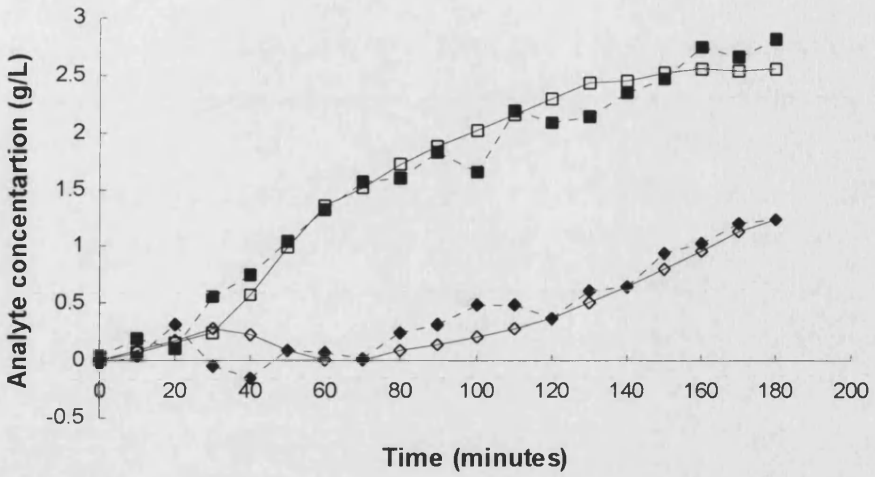


Figure 4.16. 2L fed batch microbially catalysed Baeyer-Villiger bioconversion profile monitored using on-line NIR analysis.: Lactone: (■) on-line NIR, (□) gas chromatography, ketone: (♦) online NIR, (◊) gas chromatography.

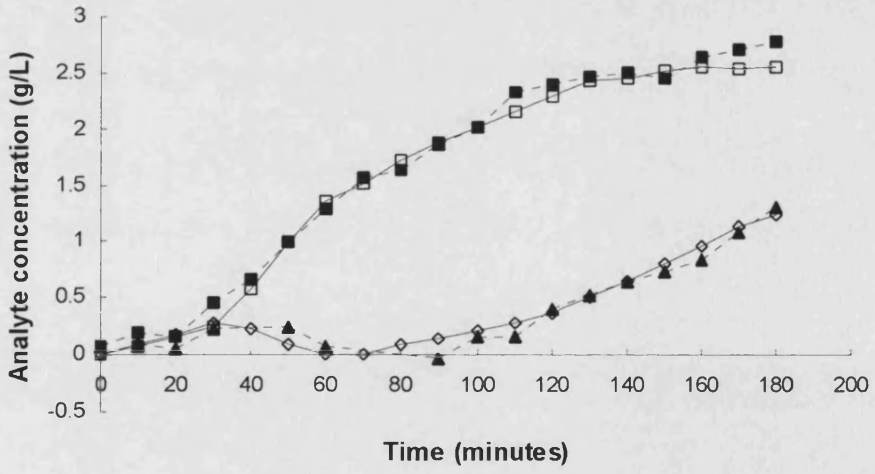


Figure 4.17. 2L fed batch microbially catalysed Baeyer-Villiger bioconversion profile monitored using at-line NIR analysis.: Lactone: (■) at-line NIR, (□) gas chromatography, ketone: (♦) at-line NIR, (◇) gas chromatography.

4.3. Discussion

4.3.1. Feasibility Study

NIR spectra of spiked study samples and process samples were taken using a transmission immersion probe. A reflectance probe was used only to analyse a commercial lyophilised sample of Lactone[1]. A reflectance NIR accessory would have been inappropriate for process samples as it depends on diffuse reflectance being absorbed by the sample and re-radiated. Some radiation is absorbed and some is scattered back in the direction of illumination. Liquid systems will affect little diffuse scatter and so no useful information can be obtained (Foss, 1998). Reflectance NIR probes are typically used for qualitative analysis of biological products (Blanco *et al.*, 1999, DeBraekeleer *et al.*, 1998, Herkert *et al.*, 2001).

The path length of the transmission cell controls the amount of NIR radiation, which is absorbed by a given matrix (sample) and therefore the amount of radiation reaching the instrument detector. The FOSS NIR spectrophotometer is equipped with a holographic concave diffraction grating and a lead sulfide (PbS) reflectance detector module. A matrix that absorbs strongly NIR radiation can result in no radiation imparting on the detector; the detector is said to flood. This results in loss of analytical information from all or part of the ensuing spectra. The water peak of 1850-2050nm in Figure 4.1 shows the detector flooding as the peak is shown to be flat. The correct path-length for the transmission cell should limit detector flooding whilst maintaining the absorber of interest below two absorbance units (S.Halsey, FOSS, personnel communication). To achieve a good response in the wavelength region of interest it was necessary to flood some parts of the spectrum in order to get extra sensitivity for the absorbance in the region of interest. With this in mind, a transmission path-length of 1.5mm was fixed for all at-line and on-line sampling.

The wavelength region of interest for quantitative NIR analysis of the analytes of interest was found to be 1550 to 1850 nm; the 1st overtone region of absorption. The NIR spectrophotometer was demonstrated to respond

linearly in the concentration range of interest for lactone (Figures 4.4), and ketone (Figures 4.5). The collected sample spectra was referenced to water, not air, in this example so that any observable absorption differences due to analyte concentration within the spectral data set could be clearly identified. There was also found to be significant contribution by glycerol within this spectral region (Figure 4.3). The absorption of NIR radiation by glycerol meant that even in the simplest biotransformation medium; water, there were at least four absorbers in addition to any cellular biocatalyst components. Water as a biotransformation medium does not allow the reaction to proceed (section 3.2.2) unlike the more complex medium type. The undefined complex medium type contains many components many of which absorb NIR radiation in the wavelength region of interest. Since the contribution to intensity of wavelength variables is additive and the biotransformation matrix complex, a multivariate approach to quantitative calibration was chosen (section 4.1.2).

At-line samples were found to be more reproducible than on-line samples (Figure 4.8) but had the disadvantage of a longer sampling time (Figure 4.7). The absence of cells, agitation and aeration from the at-line sample removes a source of radiation scattering. Scattering of NIR radiation causes base-line shifts, which is attributable to the reduced reproducibility of on-line scans compared to at-line scans (Figure 4.9). Analysis of the sensitivity of the NIR instrument in both the at-line and on-line configuration was important since the sampling error is carried over into any multivariate calibration models.

The bioconversion medium included undefined medium of the preceding *E.coli* Top10 [pQR 239] fermentation with the batch addition of glycerol and fed-batch addition of ketone (section 3.2.2). Such a bioprocess sample is a highly complex NIR matrix containing numerous constituents. The on-line sampling environment also included biomass and sparged air, both of which cause scattering of NIR radiation. The baseline differences between spectra result from scattering of radiation. Conversion of spectra to the second derivative reduced these effects (Hall *et al.*, 1996). As process sample spectra were presented as a ratio to air background (despite background

spectral information being predominately due to water) the second derivative calculation also served to enhance weak spectral features. A common procedure both in the application of near and mid-IR is to remove background (water) information such that spectral information attributable to the constituents of interest is not swamped by bands associated with the solvent (Osborne *et al.*, 2002). However, the spectrophotometer was more precise when a spectrum was ratioed to air rather than water. The standard deviation of intensity values between 2^{nd} derivative spectra at 10 wavelengths in the region of 1500-1700nm was 0.00559 for spectra ratioed to air and 34.43 intensity units for spectra ratioed to water (data not shown). The substantial increase in precision when spectra were ratioed to air can be explained by the fact that air is a stable background matrix compared to water.

Pollard and coworkers (Pollard *et al.*, 2001) presented their mid-IR sample spectra as a ratio to water, as this was the major solvent in the fermentation broth. They were able to achieve very good precision using the major absorber as a background. The use of air as a background for the weaker NIR absorbance allows for increased instrument sensitivity. A major requirement of the background is stability. The background should be the same every time in order to reduce the addition of noise to the spectral data. The spectrum of water as it exists in the sample is likely to be not the same as pure water, due to the change attendant upon the interactions of the water molecules with the solute. The indirect analysis of sodium chloride by NIR is an example of this. Despite sodium chloride having no absorption bands whatever in the NIR, Hirschfeld, (1985) was able to measure the salinity of seawater indirectly by the shifts induced on the water bands by the dissolved salts.

4.3.2. Quantitative Analysis

The NIR region is ideally suited to the analysis of the biotransformation broth since the weak absorption of NIR energy by organic molecules enables direct spectral acquisition without dilution or the requirement of short optical path lengths (section 1.2).

Fermentation broths are initially translucent matrices which progress to highly light scattering, opaque systems as the cells multiply. The Baeyer-Villiger biotransformation system described in section 3 occurs during the stationary phase of growth and so a gross change in the properties of the matrix due to fermentation does not occur. However the Baeyer-Villiger system utilises a viable whole-cell culture of *E.coli* Top10 [pQR239] expressing CHMO to convert the cyclic ketone to the lactone product at low concentrations. That is concentrations of ketone <1.0 gl^{-1} and combined lactone <5.0 gl^{-1} (section 3.2.2). NIR analysis of fermentation processes has involved monitoring analyte concentrations in multiple gram quantities and so instrument sensitivity is less of an issue (Riley *et al.*, 2000). A biotransformation is the selective enzymic modification of a defined pure ketone into defined final products, whereas in a fermentation process the product is typically a metabolite produced by the growing cells from a mixture or material supplied to the culture. The substrate and product(s) in the model biotransformation are similar molecules (section 4.2.1) and so a sensitive technique is required to specifically analyse the compounds present in the complex biotransformation medium.

The absence of large changes in the sample matrix is favourable for NIR monitoring, although the low concentrations of the substrate and product are not. In addition, glycerol is consumed during the biotransformation together with medium components, which are utilised during cellular metabolism. Where large changes have occurred during a fermentation process, NIR monitoring has involved segmentation of the process based on time periods so that multiple NIR calibration models can be developed (Arnold *et al.*, 2003). Such an approach has enabled monitoring even where large rheological changes occur. Temperature, pH, viscosity and other process variables fluctuate during a biotransformation process. When IR spectra are measured on-line the fluctuations influence the shape of the spectra (Wulfert *et al.*, 2001).

The complexity of the NIR spectra of the broth samples makes it difficult to determine the association between the observed spectral features and

concentration differences of the individual broth constituents (section 4.3.1). As NIR models are derived from the NIR spectra of unmodified samples, it is necessary to compensate for overlapping absorption bands, matrix effects and light scattering differences, which were all identified in the feasibility study. A variety of statistical regression techniques are available for deriving suitable spectroscopic models (section 1.2). PLS is one such technique, which can be applied to quantification of constituents whose spectral features are highly overlapped by bands of other matrix constituents. A robust PLS model will contain the fewest number of chemically relevant spectral factors, which provide a SEC that is of the same order of magnitude as the GC reference method used to derive the calibration algorithm. The number of PLS factors is critical to the success of the model since too few factors will not adequately model the spectroscopic data and too many factors will over fit the data set. This will increase the SEP calculated on unknown process data. A description of the principles used to select the appropriate number of factors is given in Appendix IX. For ketone monitored at-line, a 4 factor PLS regression model provided a correlation coefficient (R^2) of 0.955 and an SEC of 0.111 $g.l^{-1}$. The SEC achieved for ketone compares well with the standard error (SE) for the GC reference method; 0.04 $g.l^{-1}$ (section 3.1.3). The optimum SEC of 0.18 $g.l^{-1}$ for lactone at-line was achieved with 4 factors and compares to 0.07 $g.l^{-1}$ for the GC reference method. NIR spectroscopic determination of acetate, ammonium, biomass and glycerol at-line in an industrial *E.coli* fermentation achieved SEC of 1.1 to 1.3 $g.l^{-1}$ using multivariate calibration models (Hall *et al.*, 1996). This level of accuracy was sufficient for fermentation analysis since the constituents of the data-sets varied as much as 68 $g.l^{-1}$ for the biomass example. However, it is suggested (Doig *et al.*, 2003a, Doig *et al.*, 2003b) that ketone and lactone both inhibit the CHMO enzyme and so the biotransformation requires monitoring at concentrations of < 1 gl^{-1}. It was therefore important to maximise the sensitivity of the NIR transmission technique.

Validation ensured that the derived spectroscopic models (Table 4.0) were representative of the ketone and lactone analytes under investigation and not unique to the three biotransformation batch data sets used to construct the

models. So that validation data was independent of the calibration model data, a separate biotransformation data set was reserved for the validation work. The results of the validation given in Table 4.1 therefore demonstrate the ability of NIR spectroscopy to simultaneously generate immediate and accurate trend concentration data for ketone and lactone both at-line (Figure 4.17) and on-line (Figure 4.16). The measure of accuracy of prediction (SEP) achieved near parity with the GC reference method (Figure 4.12 to Figure 4.15) and this was similar to the SEC indicating optimal monitoring of analytes for this model.

The SEP is considered a reliable measure of performance of the calibration equation, but should be considered alongside the multiple correlation coefficient (R^2). The correlation coefficient indicates a measure of the lack of error of the calibration, and for all the equations, the R^2 values were greater than 0.95 indicating >95% of the variation within the data was modelled (Table 4.0). At-line models gave R^2 values closer to unity (approximately 0.98) than on-line models. On-line monitoring included biomass, agitation and aeration leading to light scattering, which produced spectral baseline offsets (Brimmer *et al.*, 1993) (section 4.2.2.4). This variation explained the higher SEP values and lower R^2 values. The robustness of the at-line system is further exemplified by the disparity between online and at-line F-values (Table 4.0). F-values give an indication of how well the accuracy of the calibration model (SEC) can be expected to hold up when unknown samples not included in the calibration are measured.

The independent validation of equations A-D (Table 4.1) by predicting unknown samples independent to the calibration sets indicated that the at-line NIR system is more accurate, as expected, at predicting values than the on-line method (Table 4.2).

The characteristics of at-line and on-line monitoring are given in section 1.2. Disadvantages of at-line monitoring include the increased response time (5 minutes) and the requirement for sampling of the reactor. A decision on the most appropriate sampling method (at-line or on-line) for NIR analysis of the Baeyer-Villiger bioconversion would therefore have to include, in addition to

control accuracy requirements, the response time required to maintain control within defined threshold values. There is an interest here to monitor the ketone and lactone at low concentrations so that strategies can be implemented to remove the effect of inhibition of the biocatalyst by the substrate. The at-line method would allow process monitoring at sub mM levels whilst on-line monitoring is less accurate but significantly quicker.

4.4. Conclusion

- This work has indicated that NIR spectroscopy can be applied to monitoring whole-cell biocatalytic reactions at-line and on-line.

- NIR has been applied to monitor simultaneously analytes at low concentration demonstrating the utility of NIR for control of bioconversion processes, which are characterised by inhibitory and or toxic substrates or products.

- At-line sampling was significantly quicker than the reference technique used to monitor the Baeyer-Villiger reaction and is accurate enough to detect mM quantities of the substrate and product.

- The ease of rapid monitoring on-line provides the potential for intelligent control strategies to be applied to bioprocesses with rapid reaction rates.

5. Bioprocess Monitoring and Control Using NIR

In this chapter, the development of the baeyer-Villiger process through the implementation of a rapid NIR monitoring and control strategy is described. The stereoselective biotransformation represents an important technology for the production of chiral chemical products and intermediates. The reaction is characterised by inhibition of the biocatalyst at useful substrate and product concentrations. Control strategies that minimise issues of biocatalyst inhibition require rapid monitoring of the substrate and products. NIR spectroscopy was used to monitor the biocatalytic process and facilitate a control strategy, which resulted in a productivity gain. The design and implementation of the monitoring and control strategy was directed by a bioprocess model that served to establish a systematic approach to development.

The NIR monitoring technique achieves rapid and simultaneous quantitative analysis of ketone and lactone concentration (section 4). The at-line approach was more accurate than on-line analysis (section 4.3.2). However, on-line analysis benefited from a significantly faster sample time. This is because unlike the at-line method there was no requirement for withdrawing and preparing samples for analysis. This chapter details a method to improve the accuracy of the on-line NIR method.

5.1. Materials and Methods

All materials and methods have been previously described in section 3.1 and 4.1. All biotransformation sample data generated for the NIR models in section 5.2 was generated using the 2L fermentation method (section 3.1.1.3.1). For section 5.3 and 5.4 the 7L *E.coli* TOP10 [pQR239] fermentation method was used (section 3.1.1.3.2). NIR sampling was by the at-line NIR method (section 4.1.1), which consisted of thirty-two co-averaged scans of each sample being ratioed to thirty two co-averaged scans of air. Multivariate PLS calibration models were constructed on the resulting data as described (section 4.1.2). To achieve accurate, precise and robust NIR

measurements the same at-line sampling method was used followed by manipulation of the spectral data to the second derivative of the absorption.

In section 5.3 and 5.4 the biocatalyst concentration in the spent broth medium was manipulated by diluting a concentrated suspension of biocatalyst with supernantant of the fermentation broth. The concentrate and supernatant were produced by microfiltration. The concentration of *E.coli* TOP10 [pQR239] in the fermentation broth was increased to between 17-19 $g.dcw.l^{-1}$ by microfiltration. A ProFlux®M12 microfiltration rig (Millipore Corporation, Bedford, MA, USA) fitted with a $0.1m^2$ Durapore PVDF cassette membrane and holder (Durapore, Pellicon 2 mini, Millipore (UK) Ltd, Watford, UK). The membrane had a molecular weight cut-off of 0.22 μm.

5.2. Results

5.2.1. NIR Directed Control of the Biotransformation

In this section, the development of the Baeyer-Villiger process through the implementation of a rapid monitoring and control strategy is described. The reaction is characterised by inhibition of the biocatalyst at useful substrate and product concentrations (section 3.2.2). Control strategies that minimise the issues of biocatalyst inhibition require rapid monitoring of the substrate and products. NIR Spectroscopy was used to monitor the biocatalytic process and facilitate a control strategy, which resulted in a 30% volumetric productivity gain. The design and implementation of the monitoring and control strategy was directed by a bioprocess model that served to establish a systematic approach to development.

In order to achieve the most robust and accurate NIR monitoring system the at-line method of sampling was used.

5.2.1.1. At-line NIR Calibration Model

At-line NIR calibration models were produced (3 in total), all containing spectra taken from 7 bioprocess batches (Table 5.0). The calibration models for combined lactone contained 158 samples taken from the 7

biotransformation processes. The calibration curve for the evaluation of ketone (Figure 5.0) has a SEC of 0.088 gl^{-1} and a correlation coefficient (R^2) of 0.89. Calibration of the lactone data (Figure 5.1) achieved SEC values of 0.098 and 0.116 gl^{-1}, which are comparable to the previous at-line model (Table 4.0). The SEC was achieved on a larger data set. This affected the F value, which at 3704 (Table 5) is significantly increased compared to the results given in Table 4.0.

Model	n	wavelength (nm)	Factors	R^2	SEC ($g.l^{-1}$)	F
Lactone (1)	158	1550-1820	3	0.988	0.098	3704
Lactone (2)	174	1550-1820	4	0.988	0.116	3169
Ketone (3)	144	1590-1800	5	0.89	0.088	226

Table 5.0: Calibration of at-line NIR spectra to gas chromatography data. All spectral data was mathematically treated; 2^{nd} derivative before applying PLS regression to build at-line equations. Seven independent Baeyer-Villige bioconversion data sets were used to build equations.

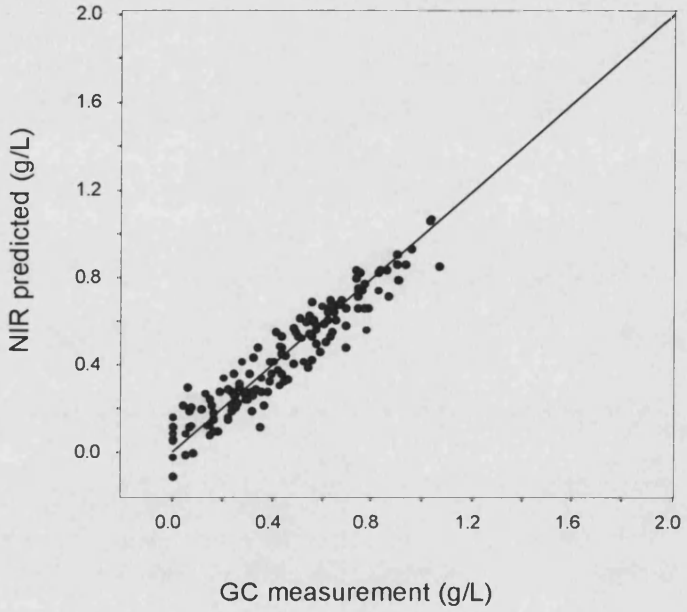


Figure 5.0. Calibration correlation plot for ketone at-line. SEC = 0.088 gl^{-1}.

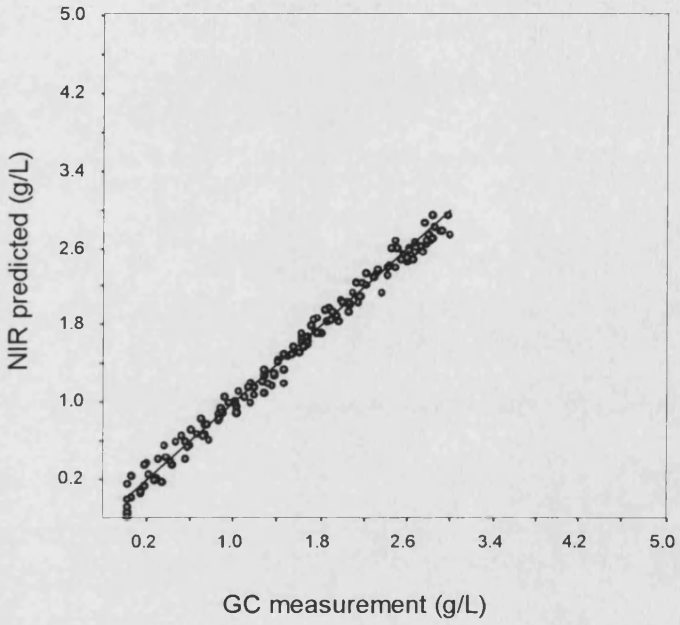


Figure 5.1. Concentration correlation plot for lactone at-line. SEC = 0.098 gl^{-1}.

5.2.1.2. Spiked At-line NIR Calibration Model

In addition to the 158 biotransformation samples, calibration model 2 (Table 5.0) contained 16 spectra taken from samples spiked with the commercial lactone[1]. This enabled the lactone concentration range to be extended from 2.8 to 4.2 gl^{-1} (Figure 5.2). Since neither of the lactone regioisomers produced during the stereoselective Baeyer-Villiger oxidation of ketone were available commercially, lactone[1] was used to spike the samples. The spiked model (model 2) had a SEC of 0.116 gl^{-1}, which was achieved using 4 PLS factors. The extra calibration factor and higher SEC compared to model 1 resulted in a lower F factor rating of 3169.

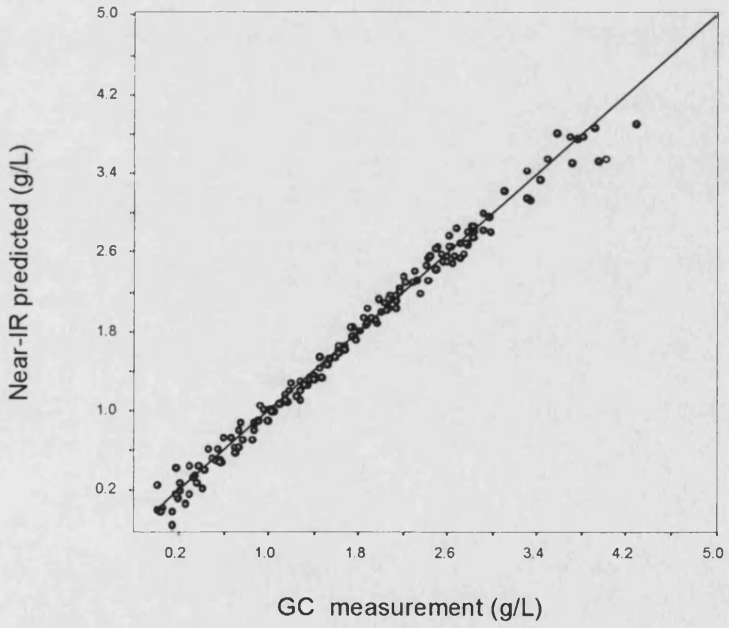


Figure 5.2. Concentration correlation plot for lactone (spiked) at-line. SEC = 0.116 gl^{-1} Lactone calibration B.

5.2.1.3. At-line NIR Monitoring and Control

Applying the calibration models (Table 5.0), the NIR spectrophotometer was used to measure simultaneously the concentrations of ketone and combined lactone at-line during a Baeyer-Villiger biotransformation experiment. The quantitative data generated using the NIR monitoring system was used to control the feed of ketone to the bioreactor to maintain the substrate concentration below inhibitory levels (<0.5 gl^{-1}). The models were later validated to establish the accuracy and robustness of the calibration models. Validation was as before (section 4.3.2), comparing the NIR data to GC reference data. The SEP is used to measure the accuracy of the NIR monitoring system. The SEP for ketone was 0.073 gl^{-1} (Figure 5.3, Table 5.1), which is similar to the SEC (Table 5.0) and so optimal for this calibration model. The results for the two lactone models differ markedly, with SEP of 0.223 and 0.118 gl^{-1} for model 1 and model 2 respectively (Figure 5.4, Table 5.1). The monitoring of the biotransformation using model 3 to quantify the ketone concentration at-line resulted in the control of the substrate below inhibitory levels. Under these conditions, a final lactone concentration of 3.6 gl^{-1} was achieved (Figure 5.5).

Model	Concentration range gl^{-1}	Number of spectra in model	SEP (gl^{-1})
Lactone (1)	0 – 2.8	158	0.223
Lactone (2)	0 – 4.0	158 + 16 spiked	0.118
Ketone (3)	0 - 1.0	144	0.073

Table 5.1 Quantitative NIR calibration models were built for prediction of reactant ketone and the product lactone during a microbially catalysed Baeyer-Villiger bioconversion. The ability of each model to predict unknown at-line samples taken from an independent biotransformation is indicated by the standard error of prediction (SEP).

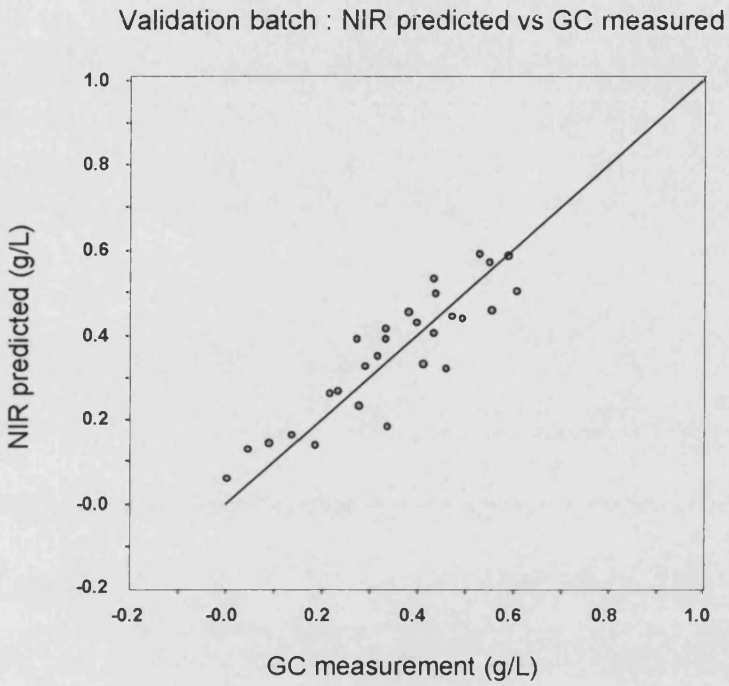


Figure 5.3. NIR calibration model validation: ketone at-line monitoring. SEP = 0.073 gl^{-1}.

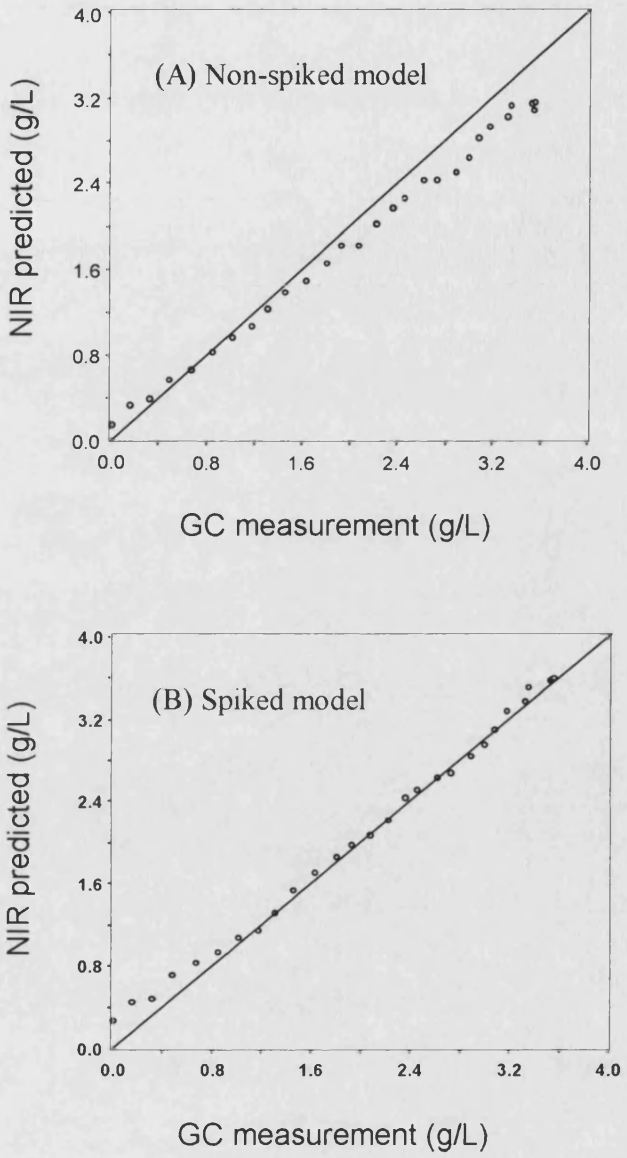


Figure 5.4. NIR calibration mocel validation, lactone at-line monitoring. SEP (model 1) = 0.223 gl^{-1}. SEP (model 2) = 0.1176 gl^{-1}.

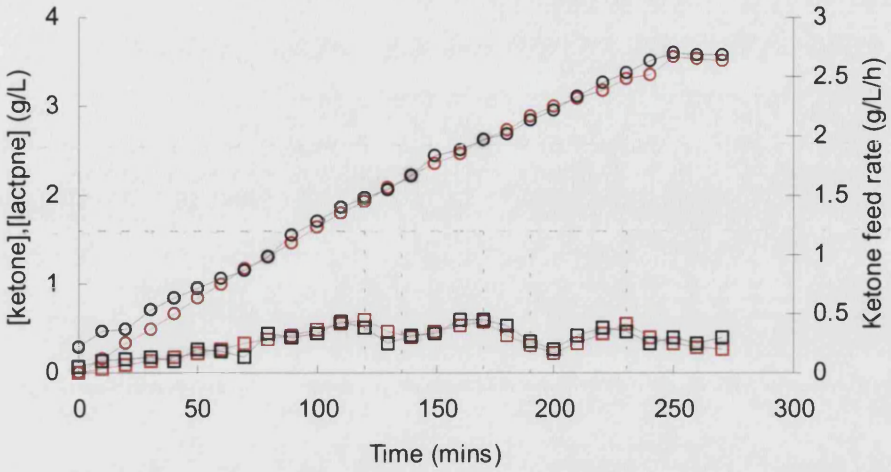


Figure 5.5. Profile of a *E.coli* Top10 [pQR 239] biotransformation process monitored and controlled using NIR spectroscopy: Ketone: (•) at-line NIR, (•) gas chromatography. Lactone: (O) at-line NIR, (O) gas chromatography. The ketone was fed into the reactor by way of a peristaltic pump, using the NIR profile to effect supervisory control.

5.2.2. On-line NIR: Membrane Covered Transmission Cells

The ability to monitor the whole-cell Baeyer-Villiger oxidation of ketone to combined lactones has been demonstrated at-line and on-line using NIR (section 4.0). The NIR monitoring technique achieves rapid and simultaneous quantitative analysis of ketone and lactone concentration. The at-line approach was more accurate than on-line analysis. However, on-line analysis benefited from a significantly faster sample time. This section details a method to improve the accuracy of the on-line NIR method.

During the development of a calibration model enough samples are included to capture the process variation (Arnold *et al.*, 2002). Where process samples are not available spiking can be used to improve calibration (section 5.2.1). During on-line calibration model development NIR sampling occurs approximately every one minute. Off-line sampling for GC analysis occurred approximately every 10 minutes. This meant only 1 in 10 NIR samples were used during calibration model development. In order to improve an NIR on-line calibration model GC reference data was interpolated increasing the available number of NIR samples for calibration. This extends the initial calibration data set leading to a more economically effective development method.

The online^{memb} technique involved sampling on-line in transmission mode. The transmission cell was covered by an absolute 0.4 µm rated membrane fabricated to fit the probe (section 5.1). The results of 7 sequential water samples taken on-line^{memb} and at-line is depicted in Figure 5.6. The on-line NIR spectra can be visually distinguished from the at-line spectra, however the 7 individual samples of each sampling method have a similar SE. The SE for at-line and online^{memb} analysis was 2.6×10^{-4} IU and 3.2×10^{-4} IU respectively. This contrasts significantly to on-line analysis (Figure 4.9) where a membrane is not used. The SE for this samples method was 8.0×10^{-4} IU.

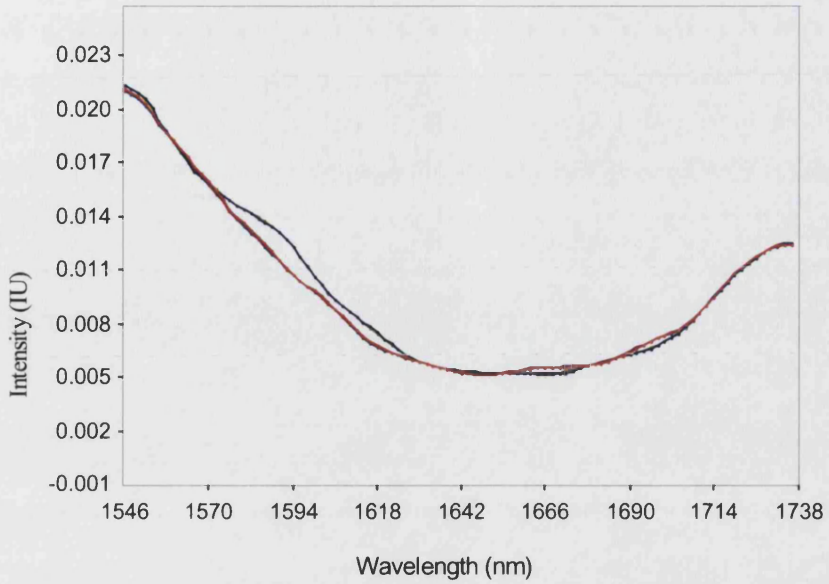


Figure 5.6. NIR 2nd derivative transmission spectra of H_20 sampled at-line (red) and on-line^{memb} over the wavelength calibration region. Intensity of the NIR spectra has arbitrary units (IU).

A whole-cell Baeyer-villiger biotransformation process was sampled using NIR online^{memb}, NIR at-line sampling and GC. The results of this single biotransformation experiment were used to generate NIR calibration algorithms using the method described in section 4.1.2.9. Unlike previous calibration models, only 1 biotransformation sample data set was acquired. Calibration models were developed on a proportion of the at-line and on-line^{memb} experimental data set for ketone (Table 5.2) and lactone (Table 5.3). The remainder of the data set was reserved for model validation. For the online^{memb} method the calibration data set was artificially expanded by interpolating the GC reference data set. The results of this interpolation are shown in Figure 5.7. This enabled calibration equations K3 and L3 to be developed using previously unusable NIR sample data. The results of the calibration experiments are given in Table 5.2 and 5.3.

Equation	Sample method	n - calibration	Factors	R^2	SEC	F
K1	At-line	28	5	0.99	0.062	309
K2	Online^{memb}	29	5	0.97	0.085	156
K3	Online^{memb} extended	135	6	0.97	0.083	584

Table 5.2: Comparison of NIR calibration methods for ketone analysis. Wavelength region for calibration, 1590 – 1800nm.

Equation	Sample method	n - calibration	Factors	R^2	SEC	F
L1	At-line	27	4	0.99	0.093	333
L2	Online^{memb}	28	5	0.98	0.091	247
L3	Online^{memb} extended	135	6	0.98	0.090	1071

Table 5.3: Comparison of NIR calibration methods for lactone analysis. Wavelength region for calibration, 1550 – 1820nm.

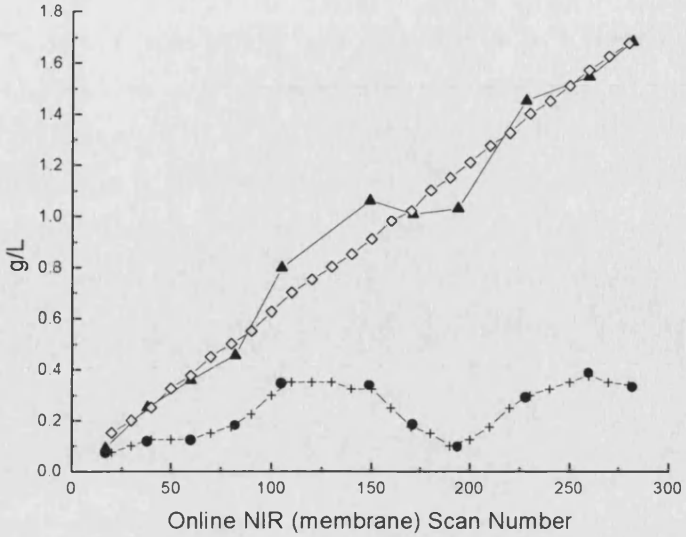


Figure 5.7. A Baeyer-villiger *E.coli* Top10 [pQR 239] biotransformation analysed by GC for ketone (•) and (▲) lactone. GC data was interpolated (ketone (+) and (◇) lactone) expanding the number of data points that could be correlated to the results of NIR online^{memb} analysis.

An internal method for validating the calibration models was used since all the sample data was the output of 1 biotransformation experiment. All 9 validation samples were removed from the model data set prior to calibration and the validation samples were chosen to be within the concentration range of ketone and lactone described by the calibration data set. These enabled predictions to be made for the purpose of validation where the data values lay within the variability of the model developed.

For both ketone and lactone analytes the SEP of validation was lowest for the at-line method; 0.06 and 0.11 gl^{-1} for ketone and lactone respectively. Prediction of validation samples online were significantly improved using equations K3 and L3 which were developed using the expanded data set. The improved SEP and correlation between GC (including interpolated data) and NIR predictions pertain a similar accuracy to that of the at-line sampling method; 0.12 and 0.15 gl^{-1} for ketone and lactone respectively.

Equation	Sample Method	n - validation	R^2	SEP
K1	At-line	9	0.99	0.062
K2	Online^{memb}	9	0.94	0.153
K3	Online^{memb} extended	9	0.98	0.118

Table 5.9: Validation of NIR calibration methods for ketone analysis. An internal validation method was used.

Equation	Sample Method	n - validation	R^2	SEP
L1	At-line	9	0.99	0.114
L2	Online^{memb}	9	0.86	0.43
L3	Online^{memb} extended	9	0.99	0.149

Table 5.10: Validation of NIR calibration methods for lactone analysis. An internal validation method was used.

5.2.3. Bioprocess Development Through Process Model Validation

Industrial implementation of biological catalytic reactions has thus far been restricted due to a variety of factors, including a lack of process engineering models. The aim of this work was to develop quantitative and predictive models to evaluate process performance, with particular emphasis on the effects of biocatalyst concentration and process design improvements which could then be used to help define process limitations and guide further experimental research.

The whole-cell CHMO catalysed Baeyer-Villiger oxidation of ketone to regio-isomeric lactones using *E.coli* Top 10 [pQR239] was modelled in a study at UCL (Doig *et al.*, 2003a). The model used process inputs of product and substrate inhibition data, oxygen mass transfer data, and stoichiometric & kinetic assumptions. The outputs of the model were product evolution rate, final yield on biomass and reaction completion time.

The modelling work was carried out as part of a research study at the Advanced Centre for Biochemical Engineering, UCL and is reproduced in part here with the permission of the authors (Doig *et al.*, 2003a).

5.2.3.1. Modelling

In the development of the process models several assumptions about the process, kinetics and stoichiometry were made:

- Typical maximum oxygen transfer rates in stirred tank fermenters are between 50 and 150 $mmol.l^{-1}.h^{-1}$, (Atkinson *et al.*, 1983) in the model a figure of 85 $mmol.l^{-1}.h^{-1}$ was used to describe the mass transfer limited volumetric reaction rate.

- For the stoichiometry of lactone production it was assumed that the experimentally determined ratio of 6.91 mole O_2 : 1 mole lactone, measured using 5 gl^{-1} biocatalyst, is valid at all other biocatalyst concentrations (Doig *et al.*, 2001).

- The inhibitory effect of product is not dependent on the biocatalyst concentration.

Based on these assumptions, the oxygen transfer rate limiting volumetric activity (R, $g.l^{-1}.min^{-1}$) was calculated according to the stoichiometry shown in Equation 5.0;

$$R = \frac{M}{1000} \times \frac{85}{6.91 \times 60} \qquad \text{(Equation 5.0)}$$

where M = 124 is the molecular weight of product. The mass transfer limiting product formation rate is 1.53 $g.l^{-1}.h^{-1}$. This rate is similar to the experimental results reported in section 3.3.2.2.

Artificial neural networks (ANNs), which can map nonlinear relationships, were used to model the specific biocatalyst activity, as a function of product and substrate concentrations. Data previously collected at 20ml and 1 litre scales (Doig *et al.*, 2001) was used. This data was normalised between 0 (min) and 1 (max) and then divided into two groups: input – product and substrate concentrations, and output – residual biocatalyst activity. The methodology used for the ANN is given in (Doig *et al.*, 2003a). The work carried out to develop the ANN was by Dr.B.Chen of UCL's Advanced centre for Biochemical Engineering. Figure 5.8 shows the comparison of the resultant simulation of the trained ANN with the experimental data.

A three term polynomial function was used to approximate the time dependent stability of the whole-cell biocatalyst. The approximation has the form of function shown in section 3.2.1.3 (Figure 3.7).

Together with the model of specific biocatalyst activity, as a function of product and substrate concentrations, the stability expression was used to predict product evolution profiles. Figure 5.9 shows the predicted biotransformation time course using biocatalyst concentrations ranging from 11 to 1790 $U.l^{-1}$ (0.2 to 32.5g $dcw.l^{-1}$). Also shown on this figure is the oxygen limiting product evolution rate of 1.53 $g.l^{-1}.h^{-1}$.

The ability of the process model to investigate a number of scenarios for the Baeyer-Villiger oxidation directed further validation bioconversions, where the biocatalyst concentration was increased after microfiltration. Validation of the model required the controlled feeding of the ketone to the reactor using near infrared spectroscopy to monitor and control the substrate concentration. This method of monitoring and control has been described previously (Bird *et al.*, 2002b).

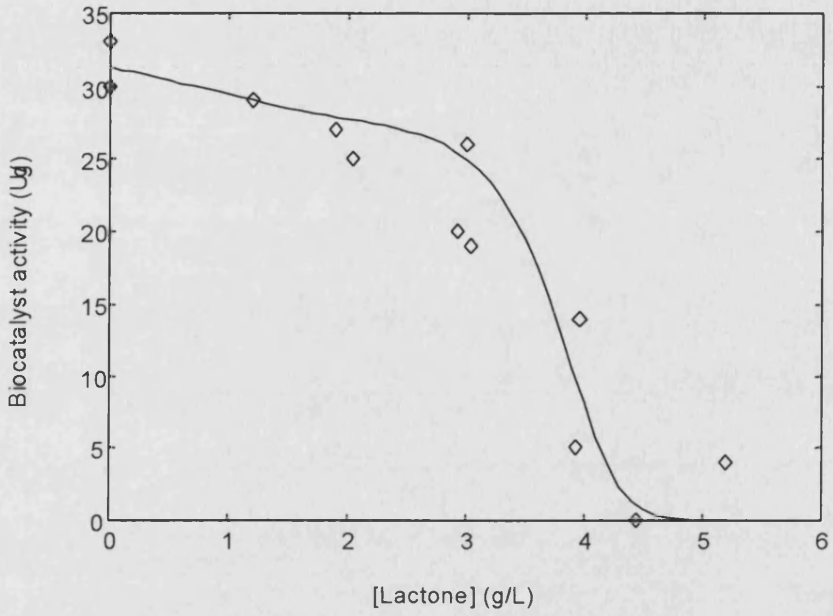


Figure 5.8 *E.coli* Top10 [pQR 239] modelled product inhibition. (◊) experimental data derived at 1 gl^{-1} ketone over range of combined lactone concentrations, (—) fitted curve. (Doig *et al.*, 2003a)

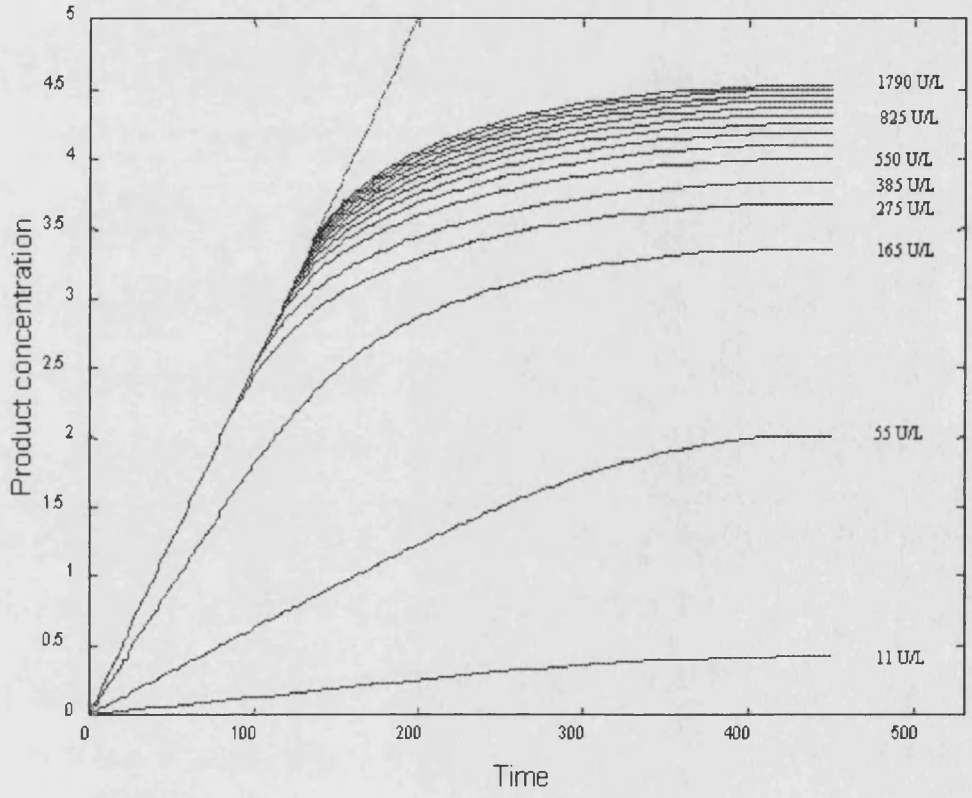


Figure 5.9 Baeyer-Villiger *E.coli* Top10 [pQR 239] biotransformation, modelled product evolution – effect of [biocatalyst] on product evolution profile and accounting for product inhibition. The tangent depicts the oxygen limiting product evolution rate of 1.53 $g.l^{-1}.h^{-1}$

5.2.3.2. Biotransformation Operating Modes

The at-line NIR technique was applied to the monitoring and control of biotransformation experiments with the purpose of testing process model assumptions given in section 5.2.3.1.

The fed-batch biotransformation can be operated in a number of ways. NIR spectroscopy was used to monitor and control the ketone concentration in the reactor, as shown in Figure 5.10. In this example, the ketone concentration is maintained below a control threshold of 1.0 gl^{-1} ketone. The formation of lactone under these conditions is linear yielding 5.1 gl^{-1} lactone on completion of the experiment at 300 minutes. The ketone concentration at 300 minutes is 0.56 gl^{-1}.

Alternatively, where the concentration of ketone was allowed to rise above the 1.0 gl^{-1} threshold, a loss of biocatalyst activity is observed (Figure 5.11A). A subsequent reduction of the ketone concentration was not matched by a recovery in the reaction rate. At 300 minutes, the lactone and ketone concentrations were 4.3 and 0.5 respectively. Stopping the ketone concentration prior to the threshold limit and allowing the remaining ketone concentration to be exhausted resulted in the biotransformation terminating at 200 minutes at a lactone and ketone concentration of 4.0 and 0 gl^{-1} respectively (Figure 5.11B).

The maximum rates of product formation for profiles A and B in Figure 5.11 are similar at 1.5 and 1.3 $g.l^{-1}.h^{-1}$ respectively. The biocatalyst material used for the experiment depicted in Figure 5.11 was the same, although profile A had double the biocatalyst concentration. This suggests that the biotransformation reaction was not being limited by a low biocatalyst concentration, and the process model (section 5.4.1.1) suggests this is due to limiting oxygen mass transfer. In contrast, the profile in Figure 5.10 has a maximum product formation rate of 1.0 $g.l^{-1}.h^{-1}$. The biocatalyst material used in this experiment (Figure 5.10) was from a different fermentation batch to profiles A and B (Figure 5.11). The lower product formation rate cannot be explained by a lower biocatalyst concentration (350 Ul^{-1}). However, it is

possible that poor expression of the CHMO gene due to process deviations in the fermentation step resulted in a lower specific activity on this occasion.

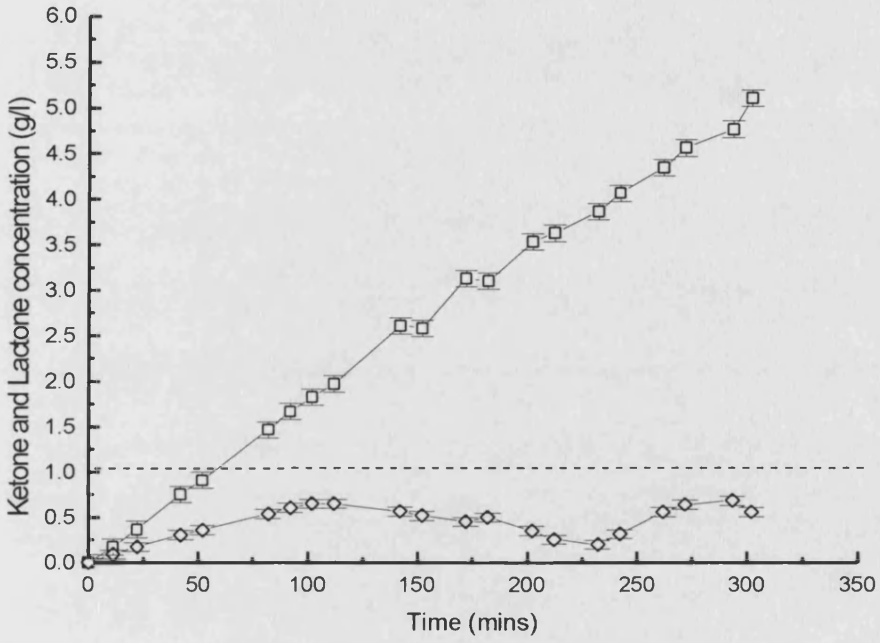


Figure 5.10.. Profile of a fed-batch *E.coli* Top10 [pQR 239] biotransformation, biocatalyst with NIR substrate feed-control. (□) Lactone. (◇) ketone. (---) Upper ketone threshold.

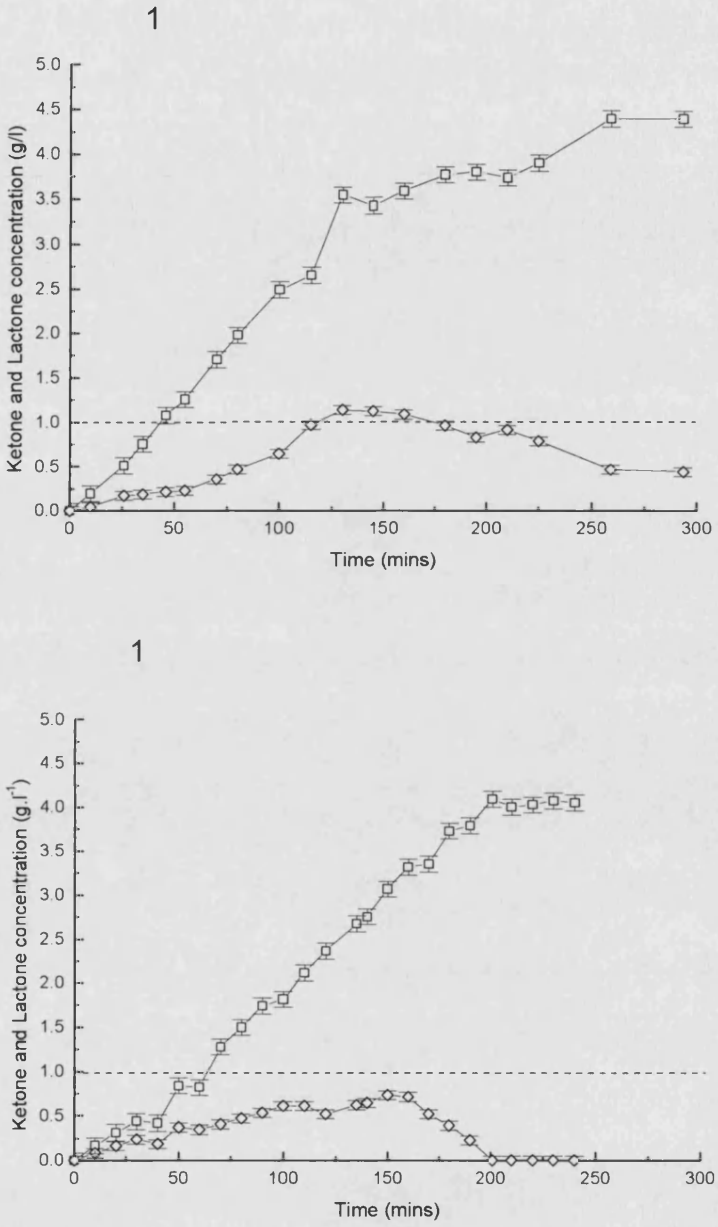


Figure 5.11. Profile of a fed-batch *E.coli* Top10 [pQR 239] biotransformations: (A) biocatalyst concentration 288 U/L where substrate concentration is allowed to reach inhibitory levels (> $1gl^{-1}$) and (B) biocatalyst concentration 519 U/L where substrate monitoring and control fails and substrate concentration decreases to 0 gl^{-1}. (□) Lactone. (◇) ketone. (----) Upper ketone threshold.

5.2.3.3. Volumetric Activity and Biocatalyst Concentration

The process model assumes the oxygen transfer rate to the reactor limits the volumetric activity. This assumption was tested by investigating the effect of changing biocatalyst concentration on the volumetric activity. So that the ketone concentration was neither exhausted or inhibiting the biocatalyst NIR at-line monitoring was used to control ketone feed rate. Seven reactions were run at biocatalyst concentrations between 1g.dcw^{-1}.l^{-1} and 20 g.dcw^{-1}.l^{-1}. The required concentrations were achieved by diluting a concentrated fermentation broth sample with spent clarified medium (section 5.1). The initial specific activity was found to be linear over a region of approximately 100 minutes for all 7 reactions. Figure 5.12 shows the lactone evolution rate of the 500 U/L (12 gl^{-1}) reaction and a reaction run at 99 U/L (1.3 gl^{-1}) at points free from substrate inhibition or limitation. The lower biocatalyst concentration reaction (1.3 g.l^{-1}) clearly has a lower initial volumetric activity, 95.2 μmol min^{-1}l^{-1} whilst the activity per gram has increased to 73.2 μmol min^{-1}g^{-1}. Plotting the initial volumetric activity against biocatalyst concentration (Figure 5.13) a maximum activity is observed, which is related to a mass transfer limiting product formation rate of 205 μmol/min/L. In the model a figure of 85 mmol.l^{-1}.h^{-1} was used to describe, the mass transfer limited volumetric reaction rate.

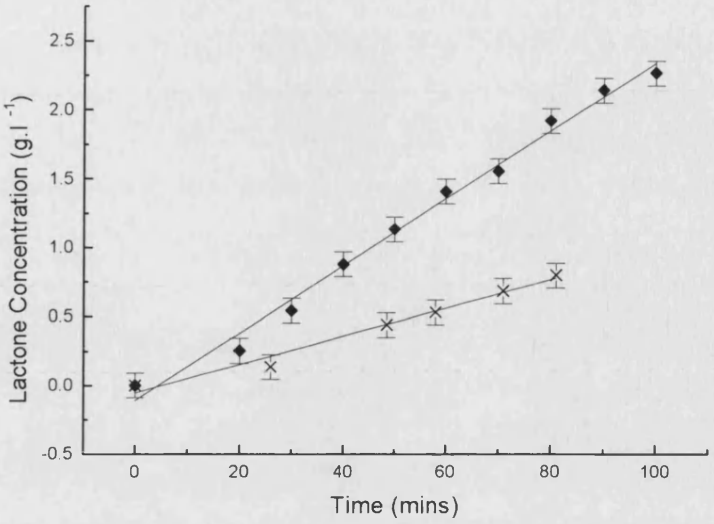


Figure 5.12 . The initial volumetric rates of activity at biocatalyst concentrations of (x) 99 UL^{-1} and 500 Ul^{-1}.

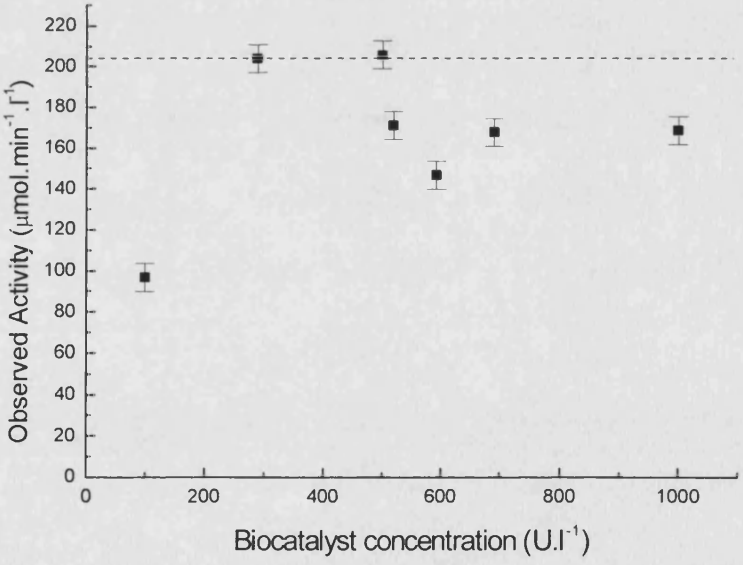


Figure 5.13. Baeyer-Villiger *E.coli* Top10 [pQR 239] biotransformations at various biocatalyst concentrations. Volumetric biocatalyst activity against biocatalyst concentration. Dashed line indicates a mass transfer limiting product formation rate of 205 μmol.min^{-1}.L^{-1}.

5.2.3.4. Specific Activity and Biocatalyst Concentration

The effect of increasing the biocatalyst concentration has a dramatic effect on the specific activity of the biocatalyst (Figure 5.14). This implicates the optimisation of the bioprocess for product on biocatalyst given that the biocatalyst has a defined stability. Stability is given as 450 minutes in the process model.

The biocatalyst concentration ($U.l^{-1}$) was calculated based on the reaction run at 1.3 gl^{-1} biocatalyst since it was found that this reaction was not limited by feed or oxygen mass transfer. The dissolved oxygen tension throughout the experiment was around 60 % rising to 80% at a point when the reaction became substrate inhibited, whereas all other reactions had a dissolved oxygen tension of zero.

The biocatalyst concentration was then calculated using this initial lactone evolution rate over its linear range and multiplied by biomass concentration for each reaction to provide the concentration in $U.l^{-1}$. This can be seen in equation 5.1 below, where R_x is the lactone evolution rate for reaction x in $\mu mol.min^{-1}.l^{-1}$, $R_{1.3}$ is the unlimited lactone evolution rate per gram of biocatalyst and is equal to 73.2 $\mu mol.min^{-1}.g^{-1}$ and $[biomasst]_x$ is the biocatalyst concentration of reaction x in $g.l^{-1}$.

Equation 5.1

$$R_x = R_{1.3g.l^{-1}} \times [biocatalyst]_x$$

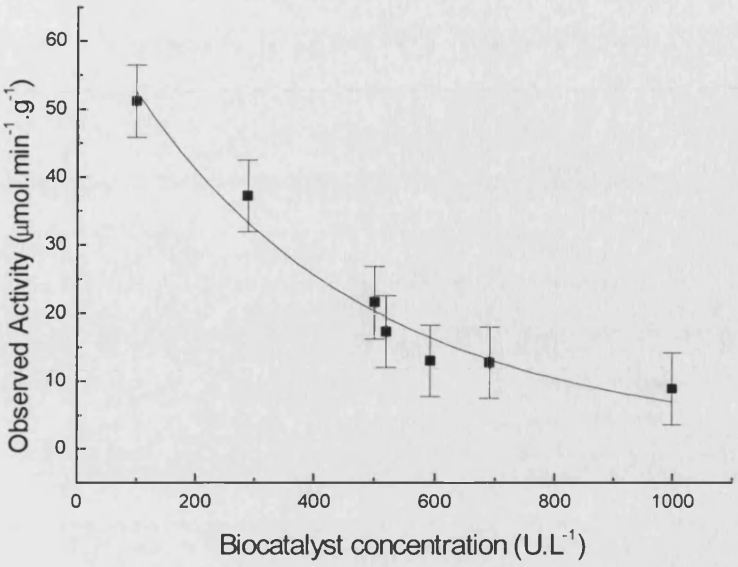


Figure 5.14. *E.coli* Top10 [pQR 239] 2L biotransformations at various biocatalyst concentrations. The specific activity of the biocatalyst depends on biocatalyst concentration.

5.2.3.5. Bioprocess Performance and Process Model Validation

The process model was validated against bioconversions controlled by NIR spectroscopy. The lactone profile process-model prediction vs experimental data for a bioconversion with 690 Ul^{-1} (13 g dcw) of biocatalyst present is shown in Figure 5.15. Figure 5.15 indicates that the model is accurate up to 3.5 gl^{-1} of combined lactone product, a point that coincides with the model's predicted drop in cellular activity i.e. 150 minutes into the experiment.(Figure 3.7) However at lactone concentrations above this point parity between the experimental and model predicted data was lost.

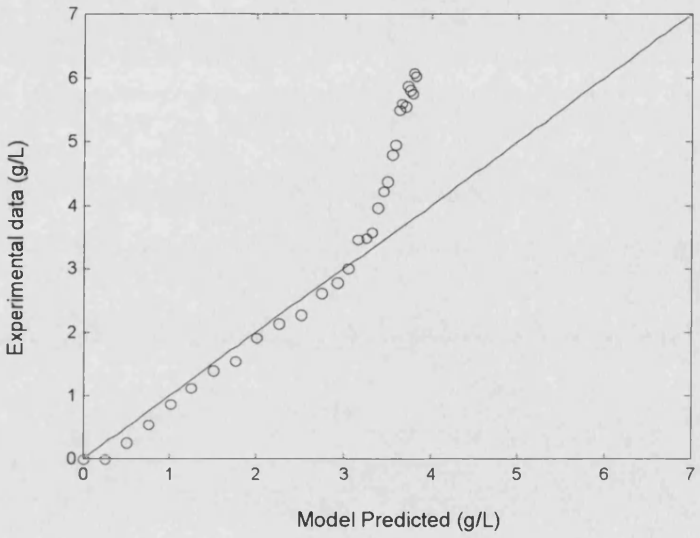


Figure 5.15. Parity of lactone profile: process-model prediction vs experimental data for a bioconversion with 690 UL^{-1} (13 g dcw) of biocatalyst present. The experimental data shown is representative of one of three bioconversions that were run at cell concentrations between 13.0 and 19.0 g $dcw.l^{-1}$.

5.3. Discussion

5.3.1. Accurate and Robust NIR Calibration

A key element in the design of any calibration model is that of accuracy and robustness. The quantitative model developed is required to contain enough of the process variability that is to be expected when running the process (Ciurezak *et al.*, 2001). The matrix for NIR sampling of the biotransformation is the broth of the preceding fermentation and so the model is required to capture batch to batch variation introduced from both the fermentation and the biotransformation steps. If spent broth medium is replaced with another medium type this variation can be controlled. However, using an alterative medium type had an adverse effect on the biocatalyst stability and activity. A difficulty in calibration model building is knowing when enough samples are taken to ensure enough of the process variation is captured.

To alleviate this problem process variation was kept to a minimum. Fermentation processes used to produce the biocatalyst were fixed, in this example (section 5.1) only biocatalyst from the LH210 fermentor was used (section 3.1.1.3.1) and the reactor was tightly controlled (section 3.2.2). Including samples from 7 Representative biotransformation experiments provided data for the PLS calibration of ketone and lactone to the GC reference data. The results (Table 5.0) generated models of sufficient accuracy for monitoring the analytes with the purpose of controlling the ketone concentration below the inhibitory level of 0.5 $g.l^{-1}$. The number of samples (and batches) required to achieve a predictive model for process control was partly intuitive. However, the calculation of the F factor from the results provided a level of confidence since this statistic (Appendix IX) indicates robustness. Calibration models containing a large numbers of samples and which demonstrate the required accuracy (SEC) with the fewest factors scores a higher F statistic. A high F statistic indicates robustness. The ketone model in Table 5.0 has a low F statistic because of the increased number of factors used in the PLS model. The models containing sample data from 7 batches (Table 5.0) were therefore considered more robust than

those containing process samples (and therefore variation) from 3 batches (Table 4.0).

The utility of the monitoring system was demonstrated during a bioconversion where samples were withdrawn every 10 minutes for analysis by NIR. Monitoring allowed the ketone concentration to be maintained below the inhibitory levels defined by the bioprocess map. Substrate feed control was achieved by intermittent pump feeding (Figure 5.5). Operation of the process with monitoring and control resulted in 3.6 gl^{-1} of lactone. This represents a 30% improvement in productivity compared to continuous fed-batch operation, where substrate inhibition curtailed the biotransformation at reduced lactone yields (Bird *et al.*, 2002b). Analysis of the process samples using the gas-chromatography reference assay confirmed the result and enabled calculation of standard errors of prediction for each NIR calibration model. The ketone SEP (Table 5.1) was sufficiently low (0.073 gl^{-1}) to allow the ketone concentration to be maintained within a 0.5 – 0.75 gl^{-1} threshold range. The spiked lactone method resulted in a significantly lower prediction error compared to the lactone model (1). Improved prediction with the spiked calibration model was due, in part, to the calibration data-set representing more closely the changes expected in the process matrix. This results in the prediction of unknowns during monitoring being interpolated and not extrapolated from the calibration model. As multivariate calibration models characterise variance, if the variability that is expected changes during bioprocess development then calibration model modifications are required.

5.3.2. Improving On-line NIR Analysis

The on-line monitoring system requires the NIR transmission probe to be permanently immersed in the biotransformation matrix within the STR. This means unlike at-line monitoring both cells (including cell matter) and sparged air are included within the sample matrix. Agitated cells and air cause scattering of the NIR radiation reducing the accuracy of the analysis (section 4.3). A filtration membrane was used to isolate the transmission probe from the cellular and aeration components of the biotransformation matrix. Using a single data set PLS regression generated calibration models for both ketone

and lactone. These models were validated using an internal but independent data set. The online^{memb} model generated was not as accurate as the at-line model. However, the robustness of the predictive model was improved by expanding the calibration data-set by interpolating the reference GC data.

The successful use of a membrane system to remove sources of noise from the sample spectra improves on-line analysis of the biotransformation. Optimising the sensitivity of the NIR method is important since the whole cell Baeyer-Villiger process operates most effectively at low concentrations of substrate and product.

5.3.3. Bioprocess Development

Using the Baeyer-Villiger process models described above it is possible to predict product - time evolution profiles using varying concentrations of biocatalyst. Figure 5.19 shows the predicted biotransformation time course using biocatalyst concentrations ranging from 11 to 1790 $U.l^{-1}$ (0.2 to 32.5g $dcw.l^{-1}$). Also shown on this figure is the oxygen limiting product evolution rate of 1.53 $g.l^{-1}.h^{-1}$. It is clear that in each case the biotransformation is completed after 450 minutes, since at this point the specific activity is zero (Figure 5.9) reflecting a loss of whole-cell viability (Figure 3.7). However, as the biocatalyst concentration increases, the biotransformation endpoint (the time point at which no further product is formed) decreases. For example, using 11 $U.l^{-1}$ (0.2g $dcw.l^{-1}$), product formation ceases after 450 minutes, whilst with 1790 $U.l^{-1}$ (32.5g $dcw.l^{-1}$), production is virtually complete after 250 minutes. These differences are due to product inhibition. At the low biocatalyst concentrations the biotransformation is limited by catalyst stability, whilst at the higher concentrations it is limited by product inhibition. This effect can also be seen from the diminishing return of final product with respect to increasing biocatalyst concentration. Nevertheless, the predicted space time yield using 1790 $U.l^{-1}$ (32.5g $dcw.l^{-1}$) biocatalyst over the 450 minute reaction time is 0.6 $g.l^{-1}.h^{-1}$ (4.8 $mmol.l^{-1}.h^{-1}$) which is comparable to values previously reported (Doig *et al.*, 2001, Doig *et al.*, 2002). The space time yield reported here (Figure 5.15) was 1.2 $g.l^{-1}.h^{-1}$ (9.6 $mmol.l^{-1}.h^{-1}$) over

260 minutes. The improvement is realised by stopping the biotransformation at 4.3 hours when the reaction rate is observed by NIR monitoring to slow.

Figure 5.15 shows the relationship between an experimental bioconversion and model predicted data for validation. The reaction shown is thought to proceed to a greater lactone concentration due to the greater control of ketone feed rate conferred by the NIR monitoring of the reaction. Parity is displayed initially, but is lost at approximately 3.5 $g.l^{-1}$, which occurs at 150 minutes into the reaction. This coincides with the process model's predicted drop in cell activity, which was influenced by previous data used to train the ANN model with respect to biocatalyst stability.

The ability to control the Baeyer-Villiger oxidation using NIR monitoring enabled the process model to be validated and model assumptions to be verified including the mass transfer limiting product formation rate of 1.53 $g.l^{-1}.h^{-1}$.

5.3.4. Conclusion

- NIR at-line monitoring was able to monitor and control a fed-batch Baeyer-Villiger biotransformation such that the substrate concentration was not under or over-fed.

- The control of the substrate feed to the bioreactor, allows ketone concentration to be controlled below inhibitory levels enabling the yield of product on substrate and biocatalyst to be optimised..

- The NIR calibration models can be extended using a spiking method where process samples are not available. The spiked NIR calibration model for the combined lactone product extended the calibration range such that validation process data was interpolated not extrapolated from the predictive model resulting in improved SEP values.

- Online NIR spectroscopy using a 0.4 μm membrane to separate the NIR transmission cell from solid phase and air bubbles in the STR was demonstrated. This novel technique achieved comparable SEC to at-line NIR calibration ketone and combined lactone model data.

- NIR monitoring and control of the whole cell Baeyer-Villiger reaction was used to validate a predictive bioprocess model which used ANN to predict time dependent product evolution profiles. The ANN process model was accurate up to 3.5 gl^{-1} of combined lactone, at a time point when the ANN model described a predicted drop in biocatalyst activity.

6. Discussion

6.1. NIR Monitoring and Control

The modelling of the Baeyer-Villiger bioprocess reaction in terms of the cumulative effect of ketone and lactone on CHMO activity has been represented graphically in a bioprocess-map (Chen *et al.*, 2002). The bioprocess map plots experimental data or data driven model results on a two-dimensional plot of one operating parameter against another (substrate concentration against product concntration). Process performance parameters e.g. biocatalyst activity, are then plotted on these axis. A process model for the studied BaeyerVilliger biotransformation model shows control of the bioprocess within an area of defined high enzyme activity (Figure 6.0) which leads to increased yields of lactone. High CHMO activity is defined as the area of the process map where enzyme activity is >70% of maximum CHMO activity after a certain time. Process-maps have particular application in controlling bioprocesses, especially where a number of time dependent variables, such as catalyst viability are of interest (Lilly *et al.*, 1996, Lye *et al.*, 2002, Woodley *et al.*, 1996).

Monitoring of ketone and lactone has traditionally been undertaken by gas chromatography, which has a response time of 35 minutes (including sampling – section 3.1.3). Such a response time negates advanced process control, especially given high CHMO activities or where there is a significant loss of enzyme activity. Loss of enzyme activity occurs when the ketone concentration is no longer maintained in the optimal concentration range (section 5.4.2). Rapid monitoring is critical in the case of the Baeyer-Villiger type oxidation. The flexibility of the control system is dependent on the time taken to determine the new values of any perturbed variable. The theoretical data in Figure 6.1 assumes the ketone concentration in the reactor to be 0.5 $g.l^{-1}$ and the upper control threshold 0.75 $g.l^{-1}$; so as to optimise CHMO activity with respect to ketone concentration (Doig *et al.*, 2003b). For a CHMO activity of 1.22 $g.l^{-1}.h^{-1}$ (10 $mmol.h^{-1}$) as depicted in Figure 3.12, there is only 6 minutes in which the control system must respond (if all the catalyst

activity were to be lost) before the ketone concentration rises above the control threshold. If the biocatalyst activity is increased to 20 $mmol.h^{-1}$ then the necessary response time drops to 3 minutes.

This work (Bird *et al.*, 2002a) indicated (section 4.3) that NIR spectroscopy can permit the sensitive and robust monitoring of analytes during a microbially catalysed Baeyer-Villiger lactone synthesis. To achieve this, multivariate calibration was employed, the purpose of which was to develop a spectral calibration model to predict the properties of interest (substrate and product concentrations). Multivariate calibration of all NIR and reference data was achieved using a partial least-square method (PLS). This factor based method (unlike multiple linear regression) does not require the user to prepare a calibration data-set where all the interfering species are known. This is especially important for bioprocesses, like the CHMO mediated Baeyer-Villiger lactone synthesis, where the considerable complexity of the reaction matrix (in particular from the spent aqueous medium of the preceding fermentation) leads to no detectable distinctive spectral peaks of analytes (section 4.2). The calibration data-set made up of samples characteristic of the process should ideally allow for all process variability and span the concentration range expected when running the process routinely. However, for CHMO mediated Baeyer-Villiger lactone synthesis, the expected (process map) lactone calibration range was not achievable using previous process samples as control of ketone concentration (which was not achievable without monitoring) resulted in an increased yield of lactone. Where expected process samples are not obtainable for the purpose of creating a spectral calibration data-set, then process samples spiked with lactone to extend the analyte concentration range can be used.

Using the bioprocess map as a development tool the calibration data set from which the NIR predictive model was originally built was suitably modified. The effect of monitoring lactone with the improved model can be seen in Figure 5.5. Operation of the process with monitoring and control resulted in 3.6 $g.l^{-1}$ of lactone, a 30% improvement in productivity over continuous fed batch operation (section 5.2) at that time. Productivity was improved further still

using this monitoring and control method (section 5.4) achieving yields of > 5.0 $g.l^{-1}$.

The often-dominant constraint of substrate and product inhibition requires careful consideration if scaleable biocatalytic reactions are to become standard options in complex syntheses. For the whole-cell Baeyer-Villiger catalysed lactone synthesis, productivity is restricted by inhibition at low concentrations. Efficient and productive operation of the process requires the controlled addition of the substrate and implementation of methods to remove the lactone product from the reactor. Such a scheme is depicted by Figure 6.2 - substrate is fed into a bioreactor in such a way that underfeeding as well as overfeeding is avoided. Likewise, lactone product is removed from the reactor using an *in situ* product removal technique (ISPR) (Lye *et al.*, 1999). Both the addition and removal of the analytes requires control of feed-back loops, which can only be realised by rapid, quantitative monitoring of both the substrate and product.

The ability to monitor rapidly requires rapid at-line or on-line monitoring. On-line monitoring is favourable since the sampling time is < 1 minute compared to 5 minutes at-line. However, the online technique when compared to at-line sampling has poorer signal reproducibility resulting in higher SEP and less robust calibration models. To overcome this, the Online^{memb} technique can be applied to bioprocess NIR systems monitoring online where multiple phases make-up the reaction matrix. This could be especially useful for the design of biotransformation processes, which seek to apply advanced control systems such as ISPR because the NIR Online^{memb} transmission probe can potentially be used rapidly and robustly to monitor a number of co-dependent operations where current probes cannot (Figure 6.2). The control of such a process is discussed in section 6.4.

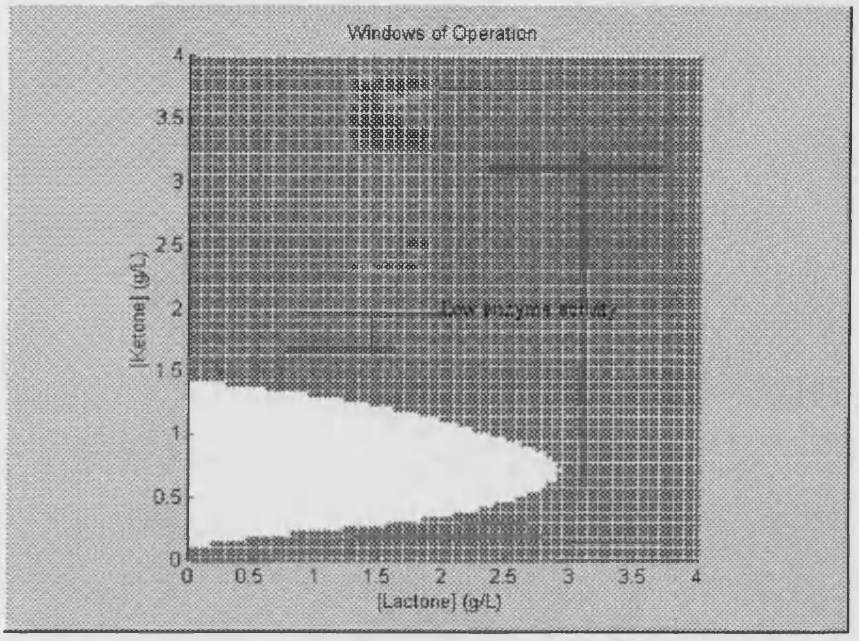


Figure 6.0. Process map for the Baeyer-Villiger bioconversion of the Baeyer-Villiger reaction in terms of the cumulative effect of ketone and lactone on CHMO activity (Chen *et al.* 2002).

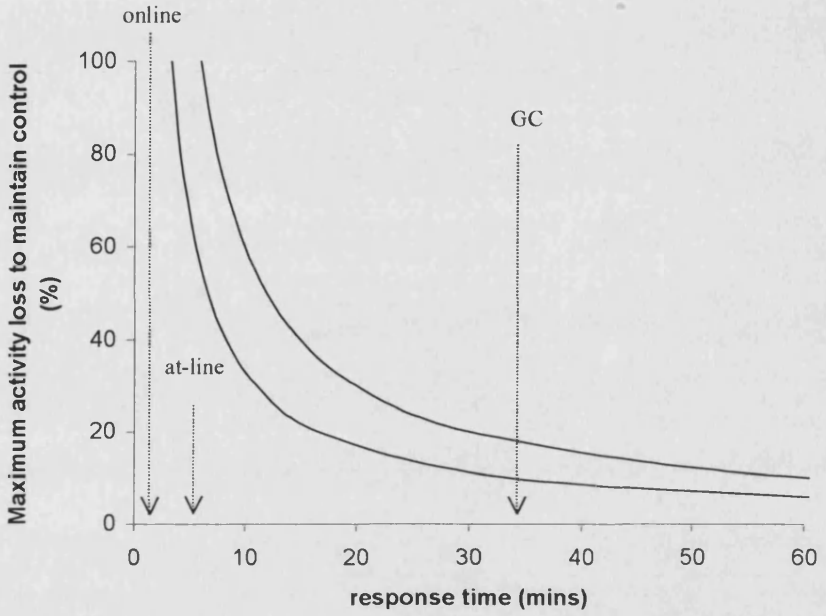


Figure 6.1. The response time required in order to implement a control action for a percentage drop in biocatalyst activity. The plot assumes a ketone concentration of 0.5 gl^{-1} and a maximum control threshold for ketone of 0.75 gl^{-1}. For a CHMO activity of 1.22 $g.l^{-1}.h^{-1}$ there is only 6 minutes in which the control system must respond. If the biocatalyst activity is increased to 2.44 $g.l^{-1}.h^{-1}$ then the necessary response time drops to 3 minutes.

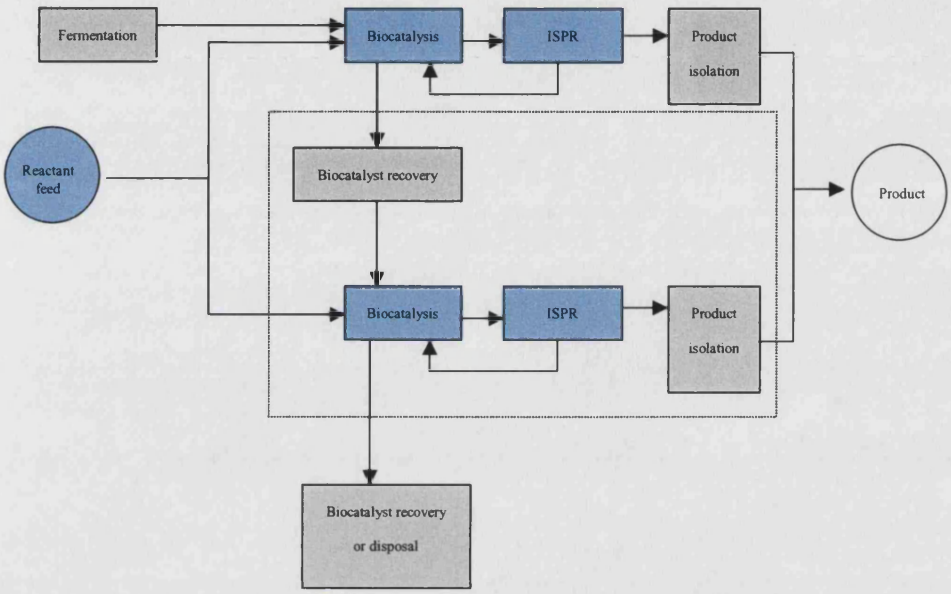


Figure 6.2. Efficient and productive operation requires the controlled addition of the substrate and an *in-situ* product removal (ISPR) technique. It is envisaged that monitoring of the substrate and product within the biocatalytic reactor could facilitate the control of the operations depicted in blue. Maintenance of the biocatalyst activity would also allow productivity improvements through biocatalyst recovery and re-use (-----).

6.2. Bioprocess Development Using NIR

A joint study at UCL's Advanced Centre of Biochemical Engineering modelled the Baeyer-Villiger bioprocess (Doig *et al.* 2003a). The models developed were used to predict – time profiles at various biocatalyst concentrations (section 5.4). The process model accounted for product inhibition, biocatalyst stability, and oxygen mass transfer limitations and they were verified by comparison to experimental data.

The ability of the process model to investigate a number of scenarios for the Baeyer-Villiger oxidation directed further validation bioconversions, where the biocatalyst concentration was increased after microfiltration (section 5.4). Validation of the model required the controlled feeding of the ketone to the reactor using NIR to monitor and control the substrate concentration.

The ketone substrate was fed periodically by adjustable peristaltic pump from an external reservoir at varying flowrates. Flowrate was controlled by the use of NIR analysis which predicted concentrations of lactone and ketone in broth samples taken at 10 minute intervals, bicycloheptanone concentration could therefore be kept at approximately 0.5 -0.75 $g.l^{-1}$ and thereby prevent substrate inhibition.

Figure 5.9 shows the predicted biotransformation time course using biocatalyst concentrations ranging from 11 to 1790 U/L. The figure indicates that the final product concentration at high biocatalyst concentration approaches 4.5 gl^{-1} lactone. Experiments using NIR monitoring achieved lactone concentrations approaching 6 gl^{-1} (section 5.4). Validation of the process model indicated that parity with experimental data was lost at a time point of 150 minutes when the lactone concentration reached 3.5 gl^{-1} . (section 5.4.2). At this same time point the process model describes a loss of biocatalyst activity (Figure 3.7). The difference between experimental and model data is therefore likely to be due to incorrect modelling of the viability of the biocatalyst.

In this instance the process model was shown not to be completely accurate, however there is a prospect of iteratively updating the model to give a process development strategy that is outlined in Figure 6.3. The ability to control the effect of substrate inhibition throughout the period of the biotransformation resulted in biocatalyst stability and lactone inhibition limiting the yield of product on biocatalyst (rather than substrate inhibition) at the given oxygen mass transfer rate. Product inhibition could be overcome by employing an efficient ISPR technique. The use of ISPR to control the effects of lactone inhibition is discussed in (section 6.4). Alternatively, it may be possible to express the CHMO in a host less susceptible to inhibition, or to use protein engineering to lessen the inhibitory effect on the enzyme itself (Lye *et al.*, 2002).

The oxygen transfer limitation may be reduced by using oxygen enriched air, or pure oxygen itself, to increase the driving force and thus the volumetric flux. The effect of this would be to increase the mass transfer limiting volumetric activity, and thus allow a greater process efficiency at the higher biocatalyst concentrations, Improvements to the biocatalyst's stability would allow the biotransformation to occur for a greater time. In this way, the maximum yield of product on biocatalyst could be increased beyond the current limit of about 2 $g.g^{-1}$ (Doig *et al.*, 2003a).

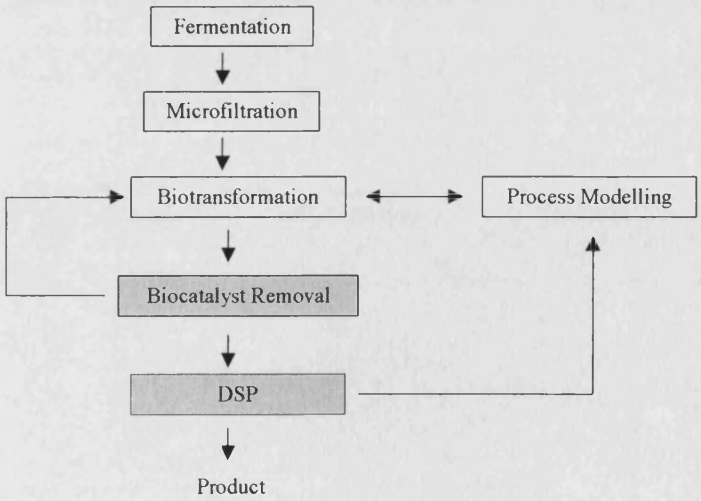


Figure 6.3.. A schematic of the whole cell BVMO bioprocess including proposed process operations (grey boxes).

6.3. Bioprocess Design Improves NIR Monitoring and Control

Where synthesis of products with multiple chiral centers is required, the use of a biotransformation step circumvents the need for multiple chemical protection and deprotection steps. The Baeyer-Villiger oxidation has only recently been achieved using transition metal chemistry and only with moderate selectivity (Peng *et al.*, 2001). Development of biotransformations with high productivity, and robust process strategies will enable biotransformations to be exploited alongside conventional chemistry to produce highly complex optically pure intermediates and products.

The capacity to monitor the biotransformation process has provided immediate advantages for process control, which result in increased productivity. However, it is recognised that the major advantages of rapid process monitoring will result from monitoring-enabled development of the bioprocess. For example, commercial bioprocess operation will depend on maximising the yield of product on the biocatalyst. To achieve this, multiple turnovers during the life of the whole-cell catalyst will be necessary. Optimal use will therefore require use of the biocatalyst under optimum kinetic conditions. Without rapid monitoring of the substrate and products to follow and control the process this cannot be achieved.

In addition to controlling ketone by feed-back control, lactone too could be controlled within the optimum activity window of the bioprocess map (Figure 6.0). This could be achieved by *in-situ* product removal (ISPR). ISPR allows the product to be removed from the vicinity of the biocatalyst during a biocatalytic reaction. Figure 6.4 schematically shows a flow sheet with an ISPR column, separate from the reactor, through which the reaction medium flows. A matrix with an affinity for the lactone allows for the removal of the product from the reaction stream so that the lactone concentration is maintained below the inhibitory concentration of 4.5 $g.l^{-1}$. Such an ISPR technique would require careful control of the flowrates through the column so as to maximise column loading and avoid the breakthrough of product back into the stirred tank reactor. Amberlite XAD4 is a suitable affinity matrix for ISPR of the lactones in the example discussed here (Simpson *et al.*,

2001). However, with this matrix the ketone binds competitively with lactone to the column. This raises the prospect of lactone yields during the product recovery stages being dependent on minimising the ketone concentration in the reaction stream. Such a holistic approach to process operation can be developed using rapid NIR monitoring during the biocatalytic step to direct the downstream processing, as shown in (Figure 6.3).

The process model has been developed (Doig *et al.*, 2003a) with product inhibition negated in the simulation. It was assumed that ISPR could be used efficiently (i.e. remove 100% of the product formed) without adversely affecting the specific and /or volumetric activity. Given this scenario the model predicts greater return of final product for increasing biocatalyst (Figure 6.5). Nevertheless, the bioconversion is still limited in time by the stability of the *E.coli* TOP10 [PQR239] biocatalyst and at high concentrations by oxygen mass transfer. Figure 6.4 depicts a Baeyer Villiger bioprocess that can be controlled within the process map (Figure 6.0).

Commercial biotransformations developed for synthesis of novel compounds require a robust and fast development strategy. Such a systematic approach to biocatalytic process design is being built (Lye *et al.*, 2002) and acceleration of such technology integration can be achieved through increased application of process modelling. As has been demonstrated in this study using NIR spectroscopy, the value of modelling to analyse process constraints can sometimes only be realised through advanced monitoring and control systems.

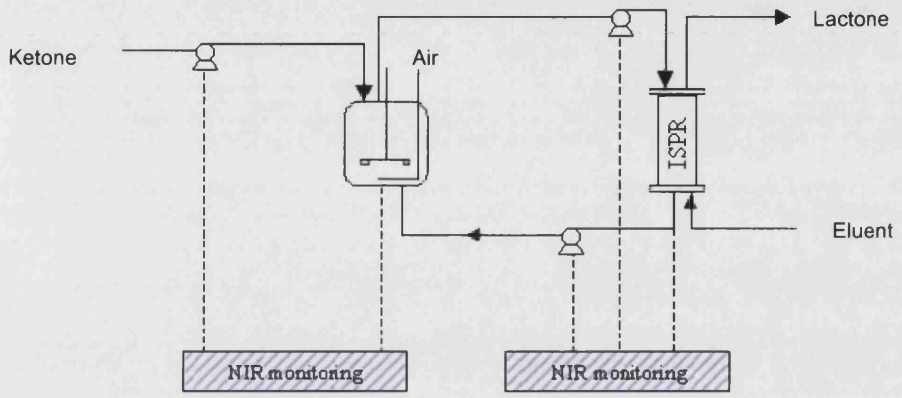


Figure 6.4. Optimum operation of the biotransformation requires the substrate and product to be monitored at or online. Monitoring by NIR spectroscopy enables feedback strategies (-----) to control ketone (feed) and lactone (in-situ product removal) to be developed.

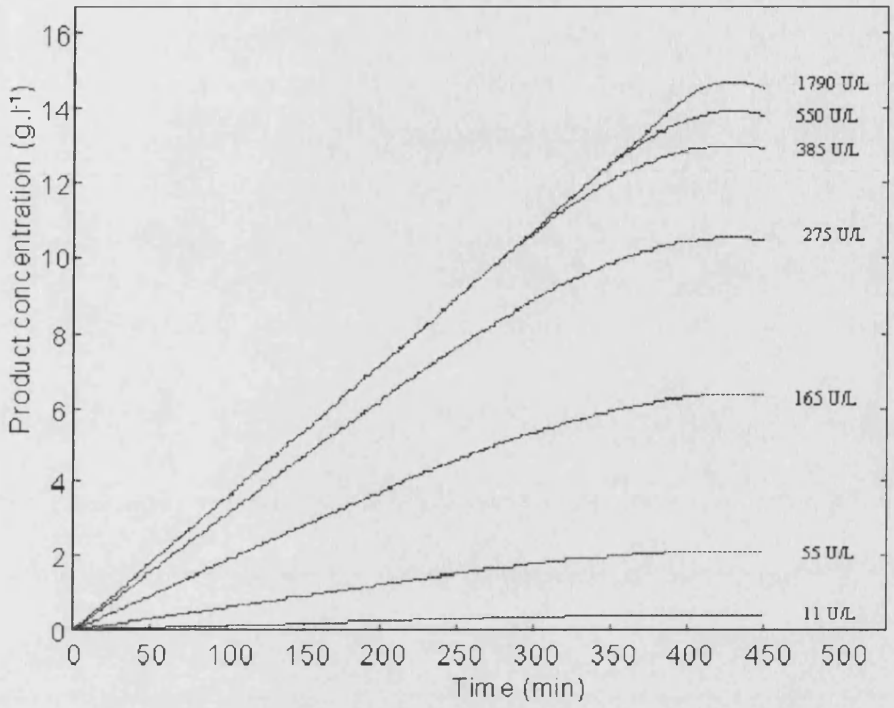


Figure 6.5.. Modelled product evolution – effect of [biocatalyst] on product evolution profile assuming product inhibition is circumvented by efficient in situ product removal.

An Innovative Enterprise Proposed Using NIR Technology

The wider use of NIR spectroscopic technology was explored as part of a study to accelerate bioprocess development in the pharmaceutical industry. The study, which was conducted at London Business School, constitutes the enterprise element of this PhD thesis. It was based on the recognition that:

- High throughput screening and the human genome project has already revolutionised the drug discovery process and enabled thousands of target compounds to be analysed compared to the traditional 50 to 100 per year for a large pharmaceutical and chemical companies.
- The chosen candidates from a given screen progress to developing a manufacturing capacity for clinical trials, and if successful, full scale production.
- To date there has been gross under-investment in manufacturing development, which now represents the bottleneck in getting a candidate to market.
- Time taken at this stage increases time to market, reduces useful patent life, delays decision-making and increases risk for investment.
- This process can take several years, even before clinical trials have started.

The RBD new technology reported is summarised below. The full report is given in appendix X.

6.3.1. Technology

Rapid Bioprocess Development (RBD) consists of accurately simulating fermentation and biotransformations at a very small scale. Using robotics one can automate thousands of these small bioreactors at the same time, and analyse the results generated using advanced instrumentation and software. The time taken in conventional optimisation (development) and

experimentation with new product candidates is largely due to the number of experiments required. Simultaneously performing these experiments allows a significant decrease in development time (equating to millions of dollars e.g. reduced time to market, extended useful patent life, higher NPV for each product).

RBD will target the biopharmaceutical and chemical bioproduct development, by bringing together three different technologies:

- Robotic liquid handling of ultra scale-down reactors,
- Advanced instrumentation for monitoring, control and analysis,
- Software for modelling and directing iterative cycles of the development.

6.3.2. Driving Force

Each year in pharmaceutical development means one more year before a patient can benefit from the drug and one less year of profit for its inventor before its patent expires and generic competitors jump in. As this patent clock winds down potential sales of well over $1 million are lost for every extra day spent in bringing the drug to market. The slowness in the development process is a result of the numerous steps that a drug must pass through en route to commercialisation. Not surprisingly all players involved in the drug development process are attacking this time expenditure.

6.3.3. Market

There are three key markets to which RBD would be targeted:

- large pharmaceutical and chemical companies
- bioprocesses consultants and contract manufacturers
- drug discovery research groups

RBD addresses the existing market of drugs and/or bio-substances development process, and in that case, RBD will replace existing, under performing practices in the drugs development industry. Drug development costs have risen to over $600 million per drug whilst 53 leading drugs will come off patent by 2005[2]. There therefore exists pressure for the pharmaceutical and chemical industry to re-asses the route to producing new-drugs.

6.3.4. Commercialisation Strategy and Recommendations

The technology to date is not sufficiently integrated for RBD to be commercialised without a development and incubation of the concept. However, the potential of RBD to deliver innovative solutions to current bioprocess development issues has led to the following recommendations:

1. Integrate and develop RBD technology
    a. Expand and define capabilities of robotically operated micro-reactors
    b. Develop a near infrared analyser for 12 well and then 96 well micro-titre plates
    c. Evaluate with UCL ventures the IPR issues surrounding RBD including the business model.
    d. Establish the demonstration of RBD as a milestone towards a spin-out business venture.
2. RBD product should be structured to deliver value to each customer segment
    a. Sell platform product to large pharmaceutical companies

b. Small to medium sized biotechnology/ chemical companies a sale can include service, partnership or license agreement.

c. RBD technology should integrate into customer development pipeline.

3. Submerge RBD into customer culture to reduce barriers to market entry

a. Paradigm shift in drug discovery programmes should be reflected in bioprocess development phase

b. Communicate value of RBD to customers in terms of man hours saved, expenditure saved, reduction in investment risk and reduction in time to market.

c. Replacement of existing expertise and resources by RBD maximise existing capabilities and not challenge them.

7. Future Work

The whole cell Baeyer-Villiger oxidation of a cyclic ketone is an important reaction of organic chemistry and demonstrates the exquisite enantioselectivity that can be achieved by application of biocatalysis. The possible commercial utility of biotransformation processes to the asymmetric oxidation of prochiral ketones is progressing (Alphand *et al.*, 2003) including the example of a dynamic resolution process to overcome the intrinsic 50% yield limitation normally imposed on a resolution process (Berezina *et al.*, 2002). Many potential commercial biotransformations including the BVMO example are currently prevented from large-scale manufacture because of a number of bottlenecks, including: lack of large scale availability of the biocatalyst; the inhibitory nature of the substrate and products; and the down stream recovery of products and substrate.

The expression of enzymes heterologously can potentially lead to recombinant organisms with improved characteristics with regard to whole-cell biocatalyst stability. This is particularly true as far as substrate and product inhibition problems are concerned. A more stable biocatalyst operating at an increased volume to produce the desired quantity of product is also required to operate at higher concentrations to minimise the expense of down stream product recovery (Straathof *et al.*, 2002). Recombinant DNA technology such as site directed mutagenesis and directed evolution could facilitate such improvements (Bornscheuer *et al.*, 2001, Burton *et al.*, 2002, May *et al.*, 2000). The use of NIR spectroscopy to monitor and control biotransformations has enabled bioreactors to be operated to control substrate feed rates below inhibitory concentrations and potentially to control the feed rate such that the substrate concentration can be optimised for down stream recovery of products.

Together with the use of process maps and models to describe and optimise bioprocess design NIR monitoring should be used to facilitate the control of biotransformations such that the maximum yield of product on biocatalyst is achieved. This is based on the prediction that recombinant whole cell

biocatalysts will continue to represent a significant proportion of a process cost (Tramper, 1996). To achieve this, re-use of the biocatalyst material should be considered. Nevertheless, a key limitation to the Baeyer-Villiger biotransformation example was product inhibition (Doig *et al.*, 2002), and so to achieve higher product concentrations in excess of 10 gl^{-1} for example, alternative process techniques such as two-liquid phase biocatalysis and ISPR should be investigated (Lye *et al.*, 1999). An ISPR technique used successfully during the biotransformation of bicycloheptanone was also found to bind substrate (Simpson *et al.*, 2001). In this case, the resin acted as a sink for product and reservoir for substrate enabling substrate concentrations as high as 20 gl^{-1}. The presence of multiple phases in bioreactors tends to complicate further NIR monitoring due to radiation scattering. This makes it difficult to distinguish analytes of interest from, gas bubbles, cells and particulate matter (Vaidyanathan *et al.*, 1999a). However, NIR online^{memb} was shown to alleviate scattering issues attributed to vessel aeration and cell matter. It should therefore be possible to use this innovative NIR^{memb} technique to monitor biotransformations in the presence of absorbent resins. There also needs to be an effort of further investigation regarding the construction of calibration data sets, including the use of sample spiking and interpolating on-line data, which should aim to improve the robustness of predictive NIR models.

Compared to the chemical route biotransformation processes can often be slow to implement whereas chemical routes, although often more expensive and less selective are better understood due to the amount of knowledge and experience available (Rozzell, 1999). To compete with chemical routes, new process development strategies are required for biotransformations. The use of automated systems at a microlitre scale to obtain key process design data has the potential to shorten development times and improve design data for biotransformations (Lye *et al.*, 2003). The pharmaceutical industry has invested heavily in the use of microplate technology for screening purposes (Rich, 2002, Stahl, 2000) and now microscale techniques are emerging as a means to develop biotransformations processes based on identified candidates. Data collected at the microscale level can be used to study

engineering operations such as equilibrium stage separation processes (Lye *et al.*, 2003). As described in Appendix X, the use of microscale technologies for quantification of process development variables is likely to be limited by the time required to assay the many samples generated using microplates. Traditional techniques, which include GC and high-pressure liquid chromatography, are not going to be applicable if the speeds of analysis are going to lag considerably the speeds of sample collection. NIR can provide sensitive and robust monitoring of complex bioprocesses on-line and rapidly (<1 minute). Currently however appropriate NIR instrumentation is only commercially available in a size designed to interface with a standard stirred tank reactor port (20mm) and not a microwell (<8mm). The process of adapting the NIR instrument to interface a modified microtiter plate is necessary. At the same time, further work on alternatives to quantitative calibration models such as PLS could address the requirement for large calibration sample data sets. For example the use of the orthogonal projection approach has been used to follow a polymorph conversion reaction to detect the end point on-line (Braekeleer *et al.*, 1999). These techniques could be the focus of future work to monitor using NIR the equilibrium parameters of biotransformations such as partition coefficients.

Conclusions

- Baeyer-Villiger monooxygenases (BVMOs) are a group of redox enzymes capable of the nucleophilic oxygenation of a wide range of linear and cyclic ketones yielding esters and lactones respectively. Cyclohexanone monooxygenase (CHMO) is the best characterised of over twenty BVMOs that have been identified. For the purpose of this work, CHMO was expressed heterologously in *Escherichia coli* TOP10 [pQR239] (Doig *et al.*, 2002). The biotransformation characteristics for the oxidation of bicyclo[3.2.0]hept-2-en-6-one to its corresponding regio-isomeric lactones (-) 1(S), 5(R), 2-oxabicyclo[3.3.0]oct-6-en-3one (ee = 94%) and (-) 1(R), 5(S), 3-oxabicyclo[3.3.0]oct-6-en-2one (ee = 99%) has been investigated (Doig *et al.*, 2003b). The key kinetic characteristics of this strain were found to be severe substrate and product inhibition. Substrate and product inhibition of biocatalysts is a frequently observed phenomenon.

- A sensitive and robust method has been developed using near infrared spectroscopy (NIR) to monitor both at-line and on-line the reactants and product of the *Escherichia coli* whole cell Baeyer-Villiger biotransformation of a cyclic ketone to optically pure lactones. The development of a system capable of precise and rapid measurements of individual components of a complex bioprocess will facilitate the realisation of improved control strategies and therefore improve productivity. Quantitative multivariate calibration of a NIR spectrophotometer for ketone and lactone resulted in a standard error of prediction at-line of 0.088 gl^{-1} and 0.110 gl^{-1} and on-line of 0.130 gl^{-1} and 0.180 gl^{-1}, respectively (Bird *et al.*, 2002a). The concentrations of substrate and product could be simultaneously monitored by NIR, which had a response time of 5.0 and 0.75 minutes at-line and on-line, respectively.

- Monitoring of ketone and lactone was previously being undertaken by a gas chromatography method, which has a response time of 35 minutes (including sampling)(Doig *et al.*, 2002). The developed NIR method satisfied the requirements for biotransformation monitoring. It was rapid, reagentless, and suitable for light scattering, heterogeneous, complex mixtures. NIR was also capable of online monitoring of the ketone substrate and combined lactone product simultaneously. The NIR spectral features of sample data were found to be broad and overlapped and so resolution of spectral data for analyte calibration to reference data required chemometric techniques for multivariate analysis.

- NIR calibration training data should include all process variability and span the concentration range expected when running the process routinely. There was found to be an issue with regard to biotransformations where substrate and product inhibition reduced productivity such that expected process samples were not obtainable for creating a spectral training data set for subsequent model development. A lactone NIR calibration model containing spiked bioprocess samples (in addition to sampled spectra) was generated. The inclusion of spiked samples improved the SEP for combined lactone from 0.22 to 0.12 gl^{-1} (Bird *et al.*, 2002b). It is considered that calibration training-sets can be designed using the spiked method to achieve a robust monitoring method for concentration ranges only feasible through advanced control [5].

- The monitoring and control of the Baeyer-Villiger bioconversion by NIR has allowed intermittent feeding of ketone such that the concentration of substrate does not rise above a threshold of 0.5 gl^{-1}. Using this feeding strategy >5.0 gl^{-1} of lactone product was produced from a single fed-batch biotransformation. This represented a 2 fold increase in productivity (Doig *et al.*, 2003a). It was apparent that without adequate monitoring and control the route to developing new biotransformation processes would be increasingly difficult.

- NIR monitoring was used to validate a process model of the Baeyer-Villiger biotransformation (Doig *et al.*, 2003a). The work indicated that the process model was accurate up to a product concentration of $3.5gl^{-1}$ after which the model underestimates combined lactone production.

- Improvements to NIR on-line monitoring were sought with the use of a membrane to separate the transmission sample cell from the adverse effects of radiation scattering caused by suspended cells, particulates and air bubbles present in the bioreactor. This innovative sampling technique achieved standard error of calibrations values similar to that of the at-line NIR technique.

1. References

Acha, V.; Meurens, M.; Naveau, H.; Agathos, S. N. (2000). ATR-FTIR sensor development for continuous on-line monitoring of chlorinated aliphatic hydrocarbons in a fixed-bed bioreactor. *Biotechnol Bioeng.* ***68***, 473-487.

Almond, M. J.; Knowles, S. J. (1999). Quantitative analysis of agrochemical formulations by multivariate spectroscopic techniques. *Appl.Spectrosc.* ***53***, 1128-1137.

Alphand, V.; Carrea, G.; Wohlgemuth, R.; Furstoss, R.; Woodley, J. M. (2003). Towards large-scale synthetic applications of Baeyer-Villiger monooxygenases. *Trends Biotechnol.* ***21***, 318-323.

Anderson, E. M.; Karin, M.; Kirk, O. (1998). One biocatalyst -many applications: the use of *Candida antarctica* B-Lipase in organic synthesis. *Biocat.Biotrans.* ***16***, 181-204.

Arnold, S. A.; Crowley, J.; Vaidyanathan, S.; Matheson, L.; Mohan, P.; Hall, J. W.; Harvey, L. M.; McNeil, B. (2000). At-line monitoring of a submerged filamentous bacterial cultivation using near-infrared spectroscopy. *Enzyme Microb.Technol.* ***27***, 691-697.

Arnold, S. A.; Harvey, L. M.; McNeil, B.; Hall, J. W. (2002). Employing near-infrared spectroscopic methods of analysis for fermentation monitoring and control: Part 1, method development. *BioPharm International.* ***November***, 26-34.

Arnold, S. A.; Harvey, L. M.; McNeil, B.; Hall, J. W. (2003). Employing near-infrared spectroscopic methods of analysis for fermentation monitoring and control: Part 1, method development. *BioPharm International.* ***January***, 47-50.

Arnold, S. A., Matheson, L., Harvey, L. M., McNeil, B. (2001). Temporarily segmented modelling: a route to improved bioprocess monitoring using near infrared spectroscopy. *Biotech.Lett.* **23**, 143 - 147.

Atkinson, B.; Mavituna, F. (1983) Biochemical engineering and biotechnology; MacMillan Publishers, UK.

Baeyer, A.; Villiger, V. (1899). Einwirkung des Caro-schen reagens auf ketone. *Ber.Dtsch.Chem.Ges.* ***32***, 3625-3633.

Balkenhohl, F.; Ditrich, K.; Hauer, B.; Ladner, W. E. (1997). Optically active amines via lipase-catalysed methoxyacetylation. *J der Praktischen Chemie.* ***339***, 381-384.

Barclay, S., Woodley, J. M., Lilly, M. D., Spargo, P. L., Pettman, A. J. (2001). Production of cyclohexanone monooxygenase from Acinetobacter calcoaceticus for large scale Baeyer-Villiger monooxygenase reactions. *Biotech.Lett.* **23**, 385 - 388.

Beauhaire, J.; Ducrot, P.; Malosse, C.; Rochat, C.; Ndiege, I. O.; Otieno, D. O. (1995). Identification and synthesis of sordidin, a male pheromone emitted by cosmopolites sordidus. *Tetrahedron Lett.* ***36***, 1043-1046.

Becker, P., Collins, A. M., Hunt, J. R., Woodley, J. M. (1997). Applications of FSQ spectrophotometric multicomponent analysis to bioconversion monitoring. *Biotechnol.Prog.* **13**, 715 - 721.

Berezina, N.; Alphand, V.; Furstoss, R. (2002). Microbiological transformations 51. The first example of a dynamic resolution process applied to a microbiological Baeyer-Villiger oxidation. *Tetrahedron Asymmetry.* ***13*** , 1953-1955.

Bird, P. A., Sharp, D. C. A., Woodley, J. M. (2002a). Near-IR spectroscopic monitoring of analytes during microbially catalysed Baeyer-Villiger bioconversions. *Organic Process Research and Development.* **6**, 569 - 576.

Bird, P. A.; Woodley, J. M.; Sharp, D. C. (2002b). Monitoring and controlling bioctalyctic processes: using biocatalysts at useful reactant and product concentrations. *BioPharm International.*, 14-21.

Blanco, M.; Coello, I.; Iturriaga, H.; Maspoch, S.; Pezuela, C.; Russo, E. (1999). Control analysis of a pharmaceutical preparation by near infrared reflectance spectroscopy. A comparative study of a spinning module and fibre optic probe. *Anal.Chim.Acta.* ***298***, 183-191.

Blanco, M., Coello, I., Montoliu, I., Romero, M. A. (2001). Orthogonal signal correction in near infrared calibration. *Anal.Chim.Acta.* **434**, 125 - 132.

Blanco, M., Coello, J., Iturriaga, H., Maspoch, S., Gonzalez-Bano, R. (2000). On-line monitoring of starch enzymatic hydrolysis by near-infrared spectroscopy. *Analyst.* **125**, 749 - 752.

Blayer, S., Dawson, M. J., Woodley, J. M., Lilly, M. D. (1996). Characterization of the chemoenzymatic synthesis of *N*-Acetyl-D-neuraminic Acid (Neu5Ac). *Biotechnol.Prog.* **12**, 758 - 763.

Blayer, S.; Woodley, J. M.; Dawson, M. J.; Lilly, M. D. (1999). Alkaline biocatalysis for the direct synthesis of N-acetyl-D-neuraminic acid (Neu5Ac) from N-acetyl-D-glucosamine (GlcNAc). *Biotechnol.Bioeng.* ***66***, 131-136.

Bommarius, A. S.; Schwarm, M.; Drauz, A. (1998). Biocatalysis to amino acid based chiral pharmaceuticals. *J.Mol Cat B - Enzymatic.* ***5***, 1-11.

Borman, S. (1990). Chirality emerges as key issue in pharmaceutical research. *Chemical and Engineering News.* ***July***, 9-14.

Bornscheuer, U. T. and Pohl, M. (2001). Improved biocatalysis by directed evolution and rational protein design. *Curr.Opin.Chem.Biol.* **5**, 137 - 143.

Boyd, D. R.; Sharma, N. D.; Allen, C. C. R. (2001). Aromatic diooxygenases: molecular biocatalysis and applications. *Curr.Opin.Biotechnol.* ***12***, 564-573.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal.Biochem.* ***72***, 248-254.

Braekeleer, K. D.; Maesschalck, R. De.; Hailey, P. A.; Sharp, D. C.; Massart, D. L. (1999). On-line application of the orthogonal projection approach (OPA) and the soft independent modelling of class analogy approach (SIMCA) for the detection of the end point of a polymorph conversion reaction by near infrared spectroscopy (NIR). *Chemometrics and Intelligent Laboratory Systems.* ***46***, 103-116.

Brimmer, P. J.; Hall, J. W. (1993). Determination of nutrient levels in a bioprocess using near-infrared spectrosocopy. *Can.J.Appl.Spectrosc.* ***38***, 155-162.

Brimmer, P. J. and Hall, J. W. (1996). Determination of nutrient levels in a bioprocess using near-infrared spectroscopy. *Can.J.Appl.Spectrosc.* **38**, 155 - 162.

Brocklebank, S., Woodley, J. M., Lilly, M. D. (1999). Immobilised transketolase for carbon-carbon bond synthesis: biocatalyst stability. *J.Mol Cat B - Enzymatic.* **7**, 223 - 231.

Bruggink, A.; Roos, E. C.; de Vroom, E. (1998). Penicillin acylase in the industrial production of beta-lactam antibiotics. *Organic Process Research and Development.* ***2***, 128-133.

Bull, A. T.; Bunch, A. W.; Robinson, G. K. (1999). Bioctalysts for clean industrial peoducts and processes. *Curr.Opin.Biotechnol.* ***2***, 246-251.

Burton, S.; Cowan, D. A.; Woodley, J. M. (2002). The search for the ideal biocatalyst. *Nature Biotechnology.* ***20***, 37-45.

Carnell, A. J.; Roberts, S. M.; Sik, V.; Willets, A. J. (1991). Microbial Oxidation of 7endo-Methylbicyclo[3.2.0]hept-2-en-6-one, 7,7-Dimethylbicyclo[3.2.0]hept-2-en-6-one and 2exo-Bromo-3endo-hydroxy-7, 7-dimethylbicyclo[3.2.0]heptan-6-one Using *Acinetobacter* NC1MB 9871. *J.Chem.Soc.Perkin.Trans.* ***10***, 2385-2389.

Cavinato, A. G.; Mayes, D. M.; Ge, Z. H.; Callis, J. B. (1990). Noninvasive method for monitoring ethanol in fermentation processes using fiber-optic near-infrared spectroscopy. *Anal.Chem.* ***62***, 1977-1982.

Chalmers, J. M. and Dent, G. (1997). Industrial analysis with vibrational spectroscopy. *RSC Analytical Spectroscopy Monographs.*,

Chauhan, R. P.; Woodley, J. M. (1997). Increasing the productivity of bioconversion processes. *Chem.Technol.* ***27***, 26-30.

Cheetham, P. (1994). Case Studies in Applied Biocatalysis - from ideas to products, in *Applied Biocatalysis*, Cabral, J. M.; Best, D.; Boross, L.; Tramper, H., editors; Harwood Academic Press, London: pp. 47-108.

Chen, B. H.; Doig, S. D.; Lye, G. J.; Woodley, J. M. (2002). Modelling of the Baeyer-Villiger monooxygenase catalysed synthesis of optically pure lactones. *Trans.I.Chem.E.(Part C).* ***80***, 51-55.

Cherug, A.; Durand, A. (1979). Optimisation of erythromycin biosynthesis by controlling pH and temperature: theoretical aspects and practical application. *Biotechnol.Bioeng.Sump.* ***9***, 303-320.

Cheymol, J. (1965). The thalidomide drama. *Rev Assoc Med Bras.* ***11***, 123-134.

Christensen, L. H.; Marcher, J.; Schulze, U.; Carlsen, M.; Min, R. W.; Nielsen, J.; Villadsen, J. (1996). Semi-on-line analysis for fast and precise monitoring of bioreaction processes. *Biotechnol.Bioeng.* ***52***, 237-247.

Chung, H. and Arnold, M. A. (1995). Monitoring the acid-catalyzed hydrolysis of starch with near-infrared spectroscopy. *Appl.Spectrosc.* **49**, 1079 - 1102.

Chung, H. and Arnold, M. A. (2000). Near-infrared spectroscopy for monitoring starch hydrolysis. *Appl.Spectrosc.* **54**, 277 - 283.

Ciurezak, E. W.; Drenen, J. K. (2001) Practical spectroscopy series 1: pharmaceutical and medical applications of near infrared spectroscopy;

Collins, A. M.; Maslin, C.; Davies, R. J. (1998). Scale-up of a chiral resolution using cross-linked enzyme crystals. *Organic Process Research and Development.* ***2***, 400-406.

Collins, A. M., Woodley, J. M., Liddell, J. M. (1995). Determination of reactor operation for the microbial hydroxylation of toluene in a two-liquid phase process. *J.Ind.Microbiol.* **14**, 382 - 388.

Corma, A.; Nemeth, L. T.; Renz, M.; Valencia, S. (2001). Sn-Zeolite Beta as a heterogeneous catalyst for Baeyer-Villiger oxidations. *Nature.* ***412***, 423-425.

Crosby, R. (1991). Synthesis of optically active compounds: A large scale perspective. *Tetrahedron Lett.*, 4789-4846.

Dadd, M. R.; Sharp, D. C.; Pettman, A. J.; Knowles, C. J. (2000). Real-time monitoring of nitrile biotransformations by mid-infrared spectroscopy. *J.Microbiol.Methods.* ***41***, 69-75.

DeBraekeleer, K.; Cuesta, S. F.; Hailey, P. A.; Sharp, D. C.; Pettman, A. J.; Massart, D. L. (1998). Influence and correction of temperature perturbations on NIR spectra during the monitoring of a polymorph conversion process prior to self- modelling mixture analysis. *J.Pharm.Biomed.Anal.* ***17***, 141-152.

Ding, Q.; Small, G. W.; Arnold, M. A. (1999). Evaluation of nonlinear model building strategies for the determination of glusose in biological matrices by near-infrared spectroscopy. *Anal.Chim.Acta.* ***384***, 333-343.

Doak, D. L.; Phillips, J. A. (1999). In situ monitoring of an *Echerichia Coli* fermentation using a diamond composition ATR probe and mid-infrared spectroscopy. *Biotechnol.Prog.* ***15***, 529-539.

Doig, S. D.; Avenell, P. J.; Bird, P. A.; Gallati,P, Koeller, K.; Lander, K. S.; Lye, J. G.; Wohlgemuth, R.; Woodley, J. M. (2002). Reactor operation and scale-up of whole cell Baeyer-Villiger catalysed lactone synthesis. *Biotechnol.Prog.* ***18***, 1039-1046.

Doig, S. D.; Bird, P. A.; Chen, B. H.; Law, H. E. M.; Woodley, J. M. (2003). Modelling whole cell biocatalytic Baeyer-Viliger oxidation processes: effect of biocatalyst concentration. *Tetrahedron Lett.*

Doig, S. D.; O'Sullivan, L. M.; Patel, S.; Ward, J. M.; Woodley, J. M. (2001). Large scale production of cyclohexanone monooxygenase from *Escherichia coli* TOP10 pQR239. *Enzyme Microb.Technol.* ***28***, 265-274.

Doig, S. D.; Simpson, H.; Alphand, V.; Furstoss, R.; Woodley, J. M. (2003). Characterisation of a recombinant *Escherichia coli* TOP10 (pQR239) whole cell biocatalyst for stereoseletive Baeyer-Villiger oxidations. *Enzyme Microb.Technol.* ***32***, 347-355.

Donoghue, N. A.; Norris, D. B.; Trudgill, P. W. (1976). The purification and properties of cyclohexanone oxygenase from *Norcardia globerula* CL1and *Acinetobacter* NC1B 9871. *Eur.J.Biochem.* ***63***, 175-192.

Donoghue, N. A.; Trudgill, P. W. (1975). The metabolism of Cyclohexanol by *Acinetabacter* NC1B9871. *Eur.J.Biochem.*, 1-7.

Doran, P. M. (1998) Bioprocess Engineering Principles; Academic Press, London.

Doyle, W.; Jennings, N. (1990). Fourier transform infrared checmical reaction monitoring using an in situ deep immeersion probe. *Spectroscopy.* ***5***, 34-38.

Ebbers, E. J.; Ariaans, G. J. A.; Houbiers, J. P. M.; Bruggink, A.; Zwaneburg, B. (1997). Controlled racemisation of optically active organic compounds: prospects for asymmetric transformations. *Tetrahedron.* ***53***, 9417-9476.

Erarslan, A.; Guray, A.; Bermek, E. (1991). Purification of kinetics of penicillin G acylase from a mutant strain of *E.coli* ATCC 11105. *J Chem Technol Biotechnol.* ***51***, 27-40.

Faber, K.; Patel, R. (2000). Chemical biotechnology. A happy marriage between chemistry and biotechnology: asymmetric synthesis via green chemistry. *Curr.Opin.Biotechnol.* ***11***, 517-519.

Federsel, H. J. (1993). Drug Chirality: scale-up, manufacturing, and control. *CHEMTECH.* ***December***

Fessenden, R. J.; Fessenden, J. S. (1982) Organic chemistry: second addition; 2nd ed.; PWS Publishers.

Fessner, W. D. (1999). Biocatalysis - from discovery to application. *Top.Curr.Chem.* ***1999***, 1-266.

Freeman, A.; Woodley, J. M.; Lilly, M. D. (1993). *In situ* product removal as a tool for bioprocessing. *Bio/Technology.* ***11***, 1007-1012.

Gagnon, R.; Grogan, G.; Levitt, M. S.; Roberts, S. M.; Wan, P. W. H.; Willets, A. J. (1994). Biological Baeyer-Villiger oxidation of some moncyclic and bicyclic ketones using monooxygenases from *Acinetobacter calcoaceticus* NCIMB9871 and *Pseudomonas putida* NCIMB 10007. *Trans 1.* 2537-2543.

Galindo, E.; Lagunas, F.; Osuna, J.; Soberon, X.; Garcia, J. L. (1998). A microbial biosensor for 6-aminopeniciloanic acid. *Enzyme Microb.Technol.* ***23***, 331-334.

Ge, Z. H., Cavinato, A. G., Callis, J. B. (1994). Noninvasive spectroscopy for monitoring cell density in a fermentation process. *Anal.Chem.* **66**, 1354 - 1362.

González-Vara, Y. R. A.; Vaccari, G.; Dosi, E.; Trilli, A.; Rossi, M.; Matteuzzi, D. (2000). Enhanced production of L-(+)-lactic acid in chemostat by Lactobacillus casei DSM 20011 using ion-exchange resins and cross-flow filtration in a fully automated pilot plant controlled via NIR. *Biotechnol.Bioeng.* ***67***, 147-156.

Gradley, M. L. and Knowles, C. J. (1994). Asymmetric hydrolysis of chiral nitriles by *Rhodococcus rhodochrous*. *Biotech.Lett.* **16**, 41 - 46.

Graham, D.; Pereira, R.; Barfield, D.; Cowan, D. (2000). Nitrile biotransformations using free and immobilized cells of a thermophilic Bacillus spp. *Enzyme Microb.Technol.* ***26***, 368-373.

Grifantini, R.; Galli, G.; Carpani, G.; Pratesi, C.; Frascotti, G.; Grandi, G. (1998). Efficient conversion of 5-subsitituted hydantoins to D-amino acids using recombinant *Escherichia coli* strains. *Microbiology.* ***144***, 947-954.

Griffin, D. R.; Gainer, J. L.; Carta, G. (2001). Asymmetric ketone reduction with immobilized yeast in hexane: biocatalyst deactivation and regeneration. *Biotechnol Prog.* ***17***, 304-310.

Guengerich, F. P. (2001). Uncommon P450-catalysed reactions. *Curr.Drug Metab.* ***2***, 93-115.

Hack, C. J., Woodley, J. M., Lilly, M. D., Liddell, J. M. (2000). Design of a control system for biotransformation of toxic substrates: toluene hydroxylation by *Pseudomonas putida* UV4. *Enzyme Microb.Technol.* **26**, 530 - 536.

Hagman, A.; Sivertsson, P. (1998). The use of NIR spectrosocopy in monitoring and controlling bioprocesses. *Process Cont Qual.* ***11***, 125-128.

Hall, J. W., McNeil, B., Rollons, M. J., Drapper, I., Thompson, B. G., Macaloney, G. (1996). Near-Infrared Spectroscopic Determination of Acetate, Ammonium, Biomass, and Glycerol in an Industrial *Escherichia coli* Fermentation. *Appl.Spectrosc.* **50**, 102 - 108.

Harper, D. (1983). *Manuf.Chemist and Aerosol News.*, 304

Held, H.; Schmid, A.; Kohler, H. E.; Suske, W.; Witholt, B.; Wubbolts, M. G. (1999). An integrated process for the production of toxic catechols from toxic phenols based on a designer biocatalyst. *Biotechnol Bioeng.* ***62***, 641-648.

Herkert, T.; Prinz, H.; Kovar, K. (2001). One hundred percent online identity check of pharmaceutical products by near-infrared spectroscopy on the packaging line. *Eur.J.Pharm.Biopharm.* ***51***, 9-16.

Hirschfeld, T. (1983). FTIR: Now and in the future - Part 1. *European Spectroscopy News.* **51**, 13 - 21.

Hirschfeld, T. (1985). Salinity determination using NIRA. *Appl.Spectrosc.* **39**, 740 - 741.

Holland, H. L.; Weber, H. K. (2000). Enzymatic hydroxylation reactions. *Curr.Opin.Biotechnol.* ***11***, 547-553.

Howard, W. W.; Sekulic, S.; Wheeler, M. J.; Taber, G.; Urbanski, F. J.; Sistare, F. E.; Norris, T.; Aldridge, P. K. (1998). On-line determination of reaction completion in a closed loop hydrogenator using NIR spectroscopy. *Appl.Spectrosc.* ***52***, 17-22.

Jacques, J.; Collet, A.; Wilen, S. H. (1981) Enantiomers, racemates and resolutions; Wiley Interscience, New York.

Jaeger, K. E.; Reetz, M. T. (1998). Microbial lipases form versatile tools for biotechnology. *Trends Biotechnol.* ***16***, 396-403.

Karim, M. N.; Yoshida, T.; Rivera, S. L.; Saucedo, V. M.; Eikens, B.; Oh, G. (1997). Global and local neural newtwork models in biotechnology: applications to different cultivation processes . *J.Ferment.Bioeng.* ***83***, 1-11.

Kelly, D. (1998). Biotransformations, In: Biotechnology eds, in *Biotechnology*, 2nd edition ed.; Rehm, H. J.; Reed, G.; Puhler, A.; Stadler, P., editors.

Kelly, D. R. (2000). Enantioselective Baeyer-Villiger reactions. Part 1. *CHIMICA OGGI/ chemistry today.* **2nd edition**, 33 - 337.

Kelly, D. R.; Wan, P.; Tsang, J. (1998). Flavin Monooxygenases - uses as catalysts for Baeyer-Villiger ring expansion and heteroatom oxidation. In Biotechnology: biotransformation I. Ed.Kelly,D.R., *A Wiley Company.* 535-587.

Kelly, D. R.; Wright, M.; Knowles, C.; Mahdi, J. G.; Taylor, I. N. (1995). Mapping of the functional active site of Baeyer-Villigerases by substrate engineering. *J.Chem.Soc.Chem.Commun.* ***7***, 729-730.

Königsberger, K.; Griengl, H. (1994). Microbial Baeyer-Villiger reaction of Bicyclo[3.2.0]heptan-6-one-A Novel Approach to Sarkomycin A. *Bioorg.Med.Chem.* ***2***, 595-604.

Konstantinov, K. B.; Yoshida, T. (1991). Knowledge-based control of fermentation processes. *Biotechnol Bioeng.* ***39***, 479-786.

Kurtanjek, Z. (1998). Prinicipal component ANN for modelling and control of baker's yeast production. *Journal of Biotechnology.* ***65***, 23-35.

Ladner, W. E.; Ditrich, K. (1999). Biocatalytic production of chiral intermediates. *CHIMICA OGGI/ chemistry today*, 51-54.

Larsson, A.; Persson, B. A.; Backvall, J. E. (1997). Enzymatic resolution of alcohols couples with ruthenium-catalyzed racemization of the substrate. *Angew.Chem.* ***36***, 1211-1212.

Lavine, B. K. (2000). Chemometrics. *Anal.Chem.* ***72(12)***, 91-97.

Layh, N., Stolz, A., Bohme, J., Effenberger, F., Knackmuss, H. J. (2002). Enantioselective hydrolysis of a racemic naproxen nitrile and naproxen amide to S-naproxen by new bacterial isolates. *J.Biotechnol.* **33**, 175 - 182.

LeBlond, C.; Wang, J.; Larson, S.; Orella, C.; Sun, K.-Y. (1998). A combined approach to characterisation of catalystic reactions using in situ kinetic probes. *Topics in Catalysis.* ***5***, 149-158.

Leszczak, J. P.; Tran-Minh, C. (1998). Optimized enzymatic synthesis of methyl benzoate in organic medium. Operating conditions and impact of different factors on kinetics. *Biotechnol Bioeng.* ***60***, 356-361.

Levitt, M. S.; Newton, R. F.; Roberts, S. M.; Willets, A. J. (1990). Preparation of optically active 6'-flurocarbocyclic nucleosidase utilising an enantiospecific enzyme-catalysed Baeyer-Villiger type oxidation. *J.Chem.Soc.Chem.Commun.* ***8***, 619-620.

Li, K.; Frost, J. (1998). Synthesis of vanillin from glucose. *J.Am.Chem.Soc.* ***120***, 10545-10546.

Liese, A.; Filho, M. V. (1999a). Production of fine chemicals using biocatalysis. *Curr.Opin.Biotechnol.* ***10***, 595-603.

Liese, A.; Seelbach, K.; Wandrey, C. (1999b) Industrial Biotransformations; Wiley-VCH Weinheim, Cambridge.

Liese, A.; Seelbach, K.; Wandrey, C. (2000) Industrial Biotransformations; Wiley-VCH Werlag GmbH, Weinheim.

Lilly, M. D. and Woodley, J. M. (1996). A structured approach to design and operation of biotransformation processes. *J.Ind.Microbiol.* **17**, 24 - 29.

Lilly, M. D. (1994). Advances in Biotransformation Processes. *Chem.Eng.Sci.* ***49***, 151-159.

Lye, G. J.; Ayazi-Shamlou, P.; Baganz, F.; Dalby, P. A.; Woodley, J. M. (2003). Accelerated design of bioconversion processes using automated microscale processing techniques. *Trends Biotechnol.* ***21***, 29-37.

Lye, G. J.; Dalby, P. A.; Woodley, J. M. (2002). Better Biocatalytic Processes Faster: new tools for the implementation of biocatalysis in organic synthesis. *Organic Process Research and Development.* ***6***, 434-440.

Lye, G. J.; Woodley, J. M. (1999). Application of in situ product-removal techniques to biocatalytic processes . *Trends Biotechnol.* ***17***, 395-402.

Macaloney, G. (1996). Near infrared spectroscopy: the technology, its growing use in biotechnology and evaluation of its utility. *SIM news.* **46**, 129 - 132.

Macaloney, G., Hall, J. W., Rollins, M. J., Drapper, I., Anderson, K. B., Preston, J., Thompson, B. G., McNeil, B. (1997). The utility and performance of near-infrared spectroscopy in simultaneous monitoring of multiple components in a high cell density recombinant *Escherichia coli* production process. *Bioproc.Eng.* **17**, 157 - 167.

Margolin, A. (1993). Enzymes in the synthesis of chiral drugs. *Enzyme Microb.Technol.* ***15***, 266-280.

Marose, S.; Lindemann, C.; Ulber, R.; Scheper, T. (1999). Optical sensor systems for bioprocess monitoring. *Trends Biotechnol.* ***17***, 30-34.

Matsumae, H.; Furui, M.; Shibatani, T. (1993). Lipase-catalysed asymmetric hydrolysis of 3-phenylglycidic ester, the key intermediate in the synthesis of ditiazem hydrochloride. *J.Ferment.Bioeng.* ***75***, 93-98.

May, O.; Nguyen, R. T.; Arnold, F. H. (2000). Inverting enantioselctivity by directed evolution of hydantoinase for improved production of L-methionine. *Nature Biotechnology.* ***18***, 317-320.

Mazzini, C.; Lebreton, J.; Alphand, V.; Furstoss, R. (1997). A chemoenzymatic strategy for the synthesis of enantiopure (R)-(-)-Baclofen. *Tetrahedron Lett.* ***38***, 1195-1196.

McCoy, M. (1999). Biocatalysis grows for drug synthesis. *Chem.Eng.News.* ***77***, 10-14.

McShane, M. J. and Cote, G. L. (1998). Near-Infrared spectroscopy for determination of glucose, lactate and ammonia in cell culture media. *Appl.Spectrosc.* **52**, 1073 - 1078.

Michielsen, M. J.; Frielink, C.; Wijffels, R. H.; Tramper, J.; Beeftink, H. H. (2000). D-malate production by permeabilized Pseudomonas pseudoalcaligenes; optimization of conversion and biocatalyst productivity. *J.Biotechnol.* ***79***, 13-26.

Mihovilovic, M. D., Chen, G., Wang, S., Kyte, B., Rochon, F., Kayser, M. M., Stewart, J. D. (2001). Asymmetric Baeyer-Villiger oxidations of 4-mono- and 4,4-disubstituted cyclohexanones by whole cells of engineered *Escherichia coli. J.Org.Chem.* **66**, 733 - 738.

Milosevic, M.; Sting, D.; Rein, A. (1995). Diamond composite sensor for ATP spectroscopy. *Spectroscopy.* ***10***, 44-49.

Muller, G. W. (1997). Thalidomide: from tragedy to new drug discovery. *CHEMTECH.* ***27***, 21-25.

Murayama, K., Yamada, K., Tsenkova, R., Wang, Y., Ozaki, Y. (1998). Near-infrared spectra of serum albumin and r-globulin and determination of their concentrations in phosphate buffer solutions by partial least square regression. *Vib.Spectrosc.* **18**, 33 - 40.

Norris, D. B.; Trudgill, P. W. (1971). The metabolism of Cyclohexnol by *Nocardia blaberula* CL1. *Biochem.J.* ***121***, 363-370.

Ogawa, J.; Shimizu, S. (1999). Microbial enzymes: new industrial applications from tradional screening methods . *Trends Biotechnol.* ***17***, 13-21.

Osborne, B. G.; Fearn, T.; Hindle, P. H. (2002) Practical NIR spectroscopy: with applications in food and beverage analysis; Wiley-VCH Weinheim, Cambridge.

Ospina, S. (1991). Characterisation and use of a Penicillin acylase biocatalyst. *CHEMTECH.* ***53***, 205-214.

Ottolina, G.; Carrea, G.; Colonna, S.; Rückemann, A. (1996). A predictive active site model for cyclohexanone monooxygenase catalysed Baeyer-Villiger oxidations. *Tetrahedron Asymmetry.* ***7***, 1123-1136.

Ozturk, S. S.; Thrift, J. C.; Blackie, J. D.; Naveh, D. (1997). Real time monitoring and control of glucose and lactate concentrations in a mammalian cell perfusion reactor. *Biotechnol.Bioeng.* ***53***, 372-378.

Pandey, A.; Benjamin, S.; Soccoi, C. R.; Nigam, P.; Kriegner, N.; Soccoi, V. T. (1999). The realm of microbial lipases in biotechnology. *Biotechnol Appl Biochem.* ***29***, 119-131.

Parmar, A.; Kumar, H.; Marwaha, S. S.; Kennedy, J. F. (2000). Advances in enzymatic transformation of penicillins to 6-aminopenicillanic acid (6-APA). *Biotechnology Advances.* ***18***, 289-301.

Patel, R. N. (2000) Stereoselective biocatalysis; Marcel Dekker, New York, USA.

Patel, R. N. (2001). Biocatalytic synthesis of intermediates for the synthesis of chiral drug substances. *Curr.Opin.Biotechnol.* ***12***, 587-604.

Peng, Y., Feng, X., Yu, K., Li, Z., Jiang, Y., Yeung, C. (2001). Synthesis and crystal structure of bis-[(4S,5S)-4,5-dihydro-4,5-diphenyl-2-(2'-oxidophenyl-xO)oxazole-xN]copper(II) and its application in the asymmetric Baeyer-Villiger reaction. J.Or*ganomet.Chem.* **619**, 204 - 208.

Petit, F.; Furstoss, R. (1995). Synthesis of (1S, 5R)-2-8-dioxabicyclo[3.3.0]octan-3-one from its enantiomer: A subunit of clerodane derivatives. *Synthesis.* ***12***, 1517-1520.

Pollard, D. J.; Buccino, R.; Connors, C. N.; Kirschner, T. F.; Olewinski, R. C.; Saini, K.; Salmon, P. M. (2001). Real-time analyte monitoring of a fungal fermentation, at pilot scale, using in situ mid-infrared spectroscopy. *Bioprocess and Biosystems Engineering.* ***24***, 13-24.

Qing, D., Small, G. W., Arnold, M. A. (1999). Evaluation of nonlinear model building strategies for the determination of glucose in biological matrices by near-infrared spectroscopy. *Anal.Chim.Acta.* **384**, 333 - 343.

Ramirez, A.; Durand, A.; Blachere, H. T. (1981). Optimal baker's yeast production in extended fed batch culture using a computer coupled pilot-fermenter., in *Second European Congress of Biotechnology, Eastbourne*, Society of Chemical Industry, London: pp. 26.

Rani, Y. K.; Rao, R. V. S. (1999). Control of fermenters - a review. *Bioprocess Engineering.* ***21***, 77-88.

Renz, M.; Meunier, B. (1999). 100 years of Baeyer-Villiger Oxidations. *Eur.J.Org.Chem.* ***4***, 737-750.

Rich, J. O. (2002). Combinatorial biocatalysis. *Curr.Opin.Chem.Biol.* ***6***, 161-167.

Richards, A.; McCague, R. (1997). The impact of chiral technology on the pharmaceutical industry. *Chemistry and Industry.*

Riesenberg, D.; Guthke, R. (1999). High-cell density cultivation of microorganisms. *Appl.Biochem.Biotechnol.* ***51***, 422-430.

Riley, M. R.; Arnold, M. A.; Murhammer, D. W.; Walls, E. L.; DelaCruz, N. (1998). Adaptive calibration scheme for quantification of nutrients and byproducts in insect cell bioreactors by near-infrared spectroscopy. *Biotechnol.Prog.* ***14***, 527-533.

Riley, M. R. and Crider, H. M. (2000). The Effect of analyte concentration range on meaurement errors obtained by NIR spectroscopy. *Talanta.* **52**, 473 - 484.

Riley, M. R., Rhiel, M., Zhou, X., Arnold, M. A., Murhammer, D. W. (1997). Simultaneous measurement of glucose and glutamine in insect cell culture media by near infrared spectroscopy. *Biotechnol.Bioeng.* **55**, 11 - 15.

Rissom, S.; Schwarz-linek, U.; Vogel, M.; Tishkov, V.; Kragl, U. (1997). Synthesis of chiral lactones in a two-enzyme system of cyclohexanone monooxygenase and formate dehydrogenase with integrated bubble free aeration. *Tetrahedron Asymmetry.* ***8***, 2523-2526.

Roberts, S. M.; Wan, P. W. H. (1998). Enzyme-Catalysed Baeyer-Villiger Oxidations. *J.Mol Cat B - Enzymatic.* ***4***, 111-136.

Rozzell, J. D. (1999). Commercial scale biocatalysis: myths and realities. *Bioorg.Med.Chem.* ***7***, 2253-2261.

Schmid, A.; Dordick, J. S. (2001). Idustrial biocatalysis today and tommorow. *Nature.* ***409***, 258-268.

Schmidt-Dannert, C. (2001). Directed evolution of single proteins, metabolic pathways, and viruses. *Biochemistry.* **40**, 13125 - 13136.

Schnell, A. (1990) Schnell Publishing Company Inc.

Schulze, B.; Wubbolts, M. G. (1999). Biocatalysis for industrial production of fine chemicals. *Curr.Opin.Biotechnol.* ***10***, 609-615.

Schuster, C. K. (1999). Monitoring the physiological status in bioprocesses on the cellular level. *Adv.Biochem.Eng.* ***66***, 185-208.

Shanley, A. (1998). Enzymes usher in new era. *Chem.Eng.* ***21***, 63-66.

Shaw, A. D., Winson, M. K., Woodward, A. M., McGovern, A. C., Davey, H. M., Kaderbhai, N., Broadhurst, D., Gilbert, R. J., Taylor, J., Timmins, E. M., Goodacre, R., Kell, D. B. (1999). Rapid analysis of high-dimensional bioprocesses using multivariate spectroscopies and advanced chemometrics. *Adv.Biochem.Eng.*, 82 - 112.

Sheldon, M. W. (1999). Applications of oxidoreductases. *Curr.Opin.Biotechnol.* ***10***, 370-375.

Sheldon, R. (1990). Inductrial synthesis of optically active compounds. *Chemistry and Industry.* ***April***, 212-215.

Sheldon, R. A. (1996). Chirotechnology: designing economic chiral syntheses. *J.Chem.Tech.biotechnol.* ***67***, 1-14.

Sheldon, R. A. (1997). Catalysis and pollution prevention. *Chem.Ind.* **-**, 12-15.

Sheldon, R. A. (2000). Atom efficiency and catalysis in organic synthesis. *Pure Appl.Chem.* ***72***, 1233-1246.

Sheng, D.; Ballou, D. P.; Massey, V. (2001). Mechanistic studies of cyclohexanone monooxygenase: chemical properties of intermediates involved in catalysis. *Biochemistry.* ***40***, 11156-11167.

Sheridan, R.; Rein, A. (1991). FTIR system reveals secrets of chemical processes. *Res.Dev.* ***33***, 100-102.

Shewale, J. G.; Sivaraman, H. (1989). Penicillin acylase enzyme production its application in the manufacture of 6-APA. *Process Biochem.* ***24***, 146-154.

Shi, Z.; Shimizu, K. (1998). On-line metabolic pathway analysis based on metabolic signal flow diagram. *Biotechnol Bioeng.* ***58***, 139-148.

Shioya, S.; Shimizu, K.; Yoshida, T. (1999). Knowledge-baed design and operation of bioprocess systems. *Journal of Bioscience and bioengineering.* ***87***, 261-266.

Shipston, N., Lenn, M., Knowles, C. (1992). Enantioselective whole cell and isolated enzyme catalysed Baeyer-Villiger oxidations of bicyclo[3.2.0]hept-2-6-one. *J.Microbiol.Methods.* **15**, 41 - 52.

Silverstein, R. M.; Bassler, G. C.; Morill, T. C. (1981) Spectrometric identification of organic compounds; fourth ed.; John wiley & Sons, New York.

Simpson, H. D., Alphand, V., Furstoss, R. (2001). Microbiological transformations 49. Asymmetric biocatalysed Baeyer-Villiger oxidation: improvement using a recombinant *Escherichia coli* whole cell biocayalyst in the presence of an adsorbent resin. *Journal of Molecular Catalysis B: Enzymatic.* **16**, 101 - 108.

Small, G. W.; Arnold, M. A.; Marquardt, L. A. (1993). Strategies for coupling digital filtering with partial least-squares regression: application to the determination of glucose in plasma by Fourier transform near-infrared spectroscopy. *Anal.Chem.* ***65***, 3279-3289.

Sonke, T.; Kaptein, B.; Boesten, W. H. J.; Broxterman, O. B.; Kamphuis, J.; Formaggio, F.; Toniolo, C.; Rutjes, F. P. J. T.; Schoemaker, H. E. (1999). Amino acid amidase catalysed preparation and further transformations of enantiopure alpha-amino acids, in *Stereoselective Biocatalysis*, Stubenrauch, J., editor; Marcel Dekker Inc, New York: pp. 23-58.

Stahl, S. (2000). implementation of a rapid microbial screening procedure for biotransformation activities. *J.Biosci.Bioeng.* ***89***, 367-371.

Stallard, B. R. (1997). Near-IR versus Mid-IR separability of three classes of organic compounds. *Appl.Spectrosc.* **51**, 625 - 630.

Stecher, H.; Faber, K. (1997). Biocatalytic deracemization techniques: dynamic resolutions and stereoinversions. *Synthesis.* ***1***, 1-16.

Stemmer, W. P. C. (1994). Rapid evolution of a protein *in vitro* by DNA shuffling. *Nature.* **370**, 389 - 391.

Stewart, J. D. (1998). Cyclohexanone monooxygenase: A useful reagent for asymmetric Baeyer-Villiger reactions. *Curr.Org.Chem.* ***2***, 195-216.

Stewart, J. D. and Walton, A. Z. (2002). An efficient enzymatic baeyer-villiger oxidation by engineered *Escherichia coli* cells under non-growing conditions. *Biotechnol.Prog.*,

Stewart, J. D. (1997). Chemist's perspective on the use of genetically engineered microbes as re-agents for organic synthesis. *Biotechnol.Gen.Eng.Rev.* ***14***, 67-143.

Stewart, J. D. (2000). Organic transformations catalysed by engineered yeast cells and related systems. *Curr.Opin.Biotechnol.* ***11***, 363-368.

Stinson, S. C. (2000). Chiral Drugs. *Chem.Eng.News.* ***78*** , 55-78.

Straathof, A. J. J.; Panke, S.; Schmid, A. (2002). The production of fine chemicals by biotransformations. *Curr.Opin.Biotechnol.* ***13***, 548-556.

Taschner, M. J.; Chen, Q. (1991). The enzymatic Baeyer-Villiger oxidation: synthesis of the C_{11}- C_{16} subunit of Ionomycin. *Bioorg.Med.Chem.Letts* ***1***, 535-538.

Thomas, E. V. (1994). A primer on multivariate calibration. *Anal.Chem.* ***66***, 795-804.

Tosa, T.; Shibatani, T. (1995). Industrial application of immobilized biocatalysts in Japan. *Ann NY Acad Sci.* ***750***, 364-375.

Tramper, H. (1994). Case Studies in Applied Biocatalysis - from ideas to products, in *Applied Biocatalysis*, Cabral, J. M.; Best, D.; Boross, L.; Tramper, H., editors; Harwood Academic Publishers, London: pp. 47-108.

Tramper, J. (1996). Chemical versus biochemical conversion: when and how to use biocatalysts. *Biotechnol.Bioeng.* **52**, 290 - 295.

Turner, N. J. (2000). Applications of transketolases in organic synthesis. *Curr.Opin.Biotechnol.* ***11***, 527-531.

Um, P. J.; Drueckhammer, D. (1998). Dynamic enzymatic resolution of thioesters. *J.Am.Chem.Soc.* ***120***, 5605-5610.

Vaccari, G., Dosi, E., Campi, A. L., Gonzales-Vara y, R., Matteuzzi, D., Mantovani, G. (1994). A near-infrared spectroscopy technique for the control of fermentation processes: An application to lactic acid fermentation. *Biotechnol.Bioeng.* **43**, 913 - 917.

Vaidyanathan, S.; Arnold, S. A.; Matheson, L.; Mohan, P.; McNeil, B.; Harvey, L. M. (2001). Assessment of near-infrared spectral information for rapid monitoring of bioprocess quality. *Biotechnol.Bioeng.* ***74***, 376-388.

Vaidyanathan, S., Macaloney, G., McNeil, B. (1999a). Fundamental investigations on the near-infrared spectra of microbial biomass as applicable to bioprocess monitoring. *Analyst.* **124**, 157 - 162.

Vaidyanathan, S., Macaloney, G., Vaughn, J., McNeil, B., Harvey, L. M. (1999b). Monitoring of submerged bioprocesses. *Critic.Rev.Biotech.* **19**, 277 - 316.

Walton, A. Z.; Stewart, J. D. (2002). An efficient enzymatic Baeyer-Villiger oxidation by engineered *Escherichia coli* under non growing conditions. *Biotechnol Prog.* ***18***, 262-268.

Ward, O. P.; Singh, A. (2000). Enzymatic asymmetric synthesis by decarboxylases. *Curr.Opin.Biotechnol.* ***11***, 520-526.

Willets, A. J. (1997). Structutal studies and synthetic applications of Baeyer-Villiger monooxygenases. *Trends Biotechnol.* ***15***, 55-62.

Willetts, A. J. (1997). Structural studies and synthetic applications of Baeyer-Villiger monooxygenases. *Trends Biotechnol.* **15**, 55 - 62.

Woodley, J. M.; Titchener-Hooker, N. (1996). The use of windows of operation as a bioprocess design tool. *Bioproc.Eng.* ***14***, 263-268.

Workman, J.; Springsteen, A. W. (1997) Applied spectroscopy: a compact reference for practitioners; Academic Press.

Wright, M.; Knowles, C.; Petit, F.; Furstoss, R. (1994). Enantioselective inhibition studies of the cyclohexanone monooxygenase from *Acinetobacter* NCIMB 9871. *Biotech.Lett.* ***16***, 1287-1292.

Wülfert, F., Kok, W. T., Smilde, A. K. (1998). Influence of temperature on vibrational spectra and consequences for the predictive ability of multivariate models. *Anal.Chem.* **70**, 1761 - 1767.

Wynberg, H.; Staring, E. G. J. (1982). Asymmetric synthesis of (S)-and(R)-Malic Acid from Ketone and Chloral. *J.Am.Chem.Soc.* ***104***, 166

Yagasaki, M.; Ozaki, A. (1998). Industrial biotransformation for the production of D-amino acids. *J.Mol Cat B - Enzymatic.* ***4***, 1-11.

Yamada, H.; Kobayashi, M. (1996). Nitrile hydratase and its application to industrial production of Acrylamide . *Biosci.Biotechnol.Biochem.* ***60***, 1391-1400.

Yano, T., Aimi, T., Nakano, Y., Tamai, M. (1997). Prediction of the concentrations of ethanol and acetic acid in the culture broth of a rice vinegar fermentation using near-infrared spectroscopy. *J.Ferment.Bioeng.* **84**, 461 - 465.

Yano, T. and Harata, M. (1994). Prediction of the concentration of several constituents in a mouse-mouse hybridoma culture by Near Infrared spectroscopy. *Journal of Fermentation and Bioengineering.* **77**, 659 - 662.

Yeung, K. S.; Hoare, M.; Thornhill, N. F.; Williams, T.; Vaghjiani, J. D. (1999). Near-infrared spectroscopy for bioprocess monitoring and control. *Biotechnol.Bioeng.* ***63***, 684-693.

Zambianchi, F.; Pasta, P.; Carrea, G.; Colonna, S.; Gaggero, N.; Woodley, J. M. (2002). Uae of isolated cyclohexanone monooxygenase from recombinant *Escherichia Coli* as a biocatalyst for Baeyer-Villiger and sulfide oxidations. *Biotechnol Bioeng.* ***78***, 489-496.

Zambianchi, F.; Pasta, R.; Ottolina, G.; Carrea, G.; Colonna, S.; Gaggero, N.; Ward, J. M. (2001). Effect of substrate concentration on the enantioselectivity of cyclohexanone monoxygenase from *Acinetobacter caloaceticus* and its rationalisation. *Tetrahedron Asymmetry.* ***11***, 3653-3657.

Zhao, H., Giver, L., Shao, Z., Affholter, J. A., Arnold, F. H. (1998). Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nature Biotechnology.* **16**, 258 - 261.

I. Appendix: Bioreactor specification

The following is the specification for the two fermenters (LH210_a and LH2000) used to grow *E.coli* Top10 [pQR239] and the bioreactor used for the biotransformation (LH210_b).

Bioreactor	LH210_a	LH2000	LH210_b
Vessel working volume (L)	1.6	5.3	1.5
Vessel height (mm)	212	400	-
Vessel diameter (mm)	130	160	-
Aspect ratio	1.63	2.5	1.83
No. of impellers	2	2	2
Impeller diameter (mm)	40.4	62.3	-
Tip dimensions (mm)	L=10.5, W=1.0, H=7.9	L=16, W=2, H=3	-
Ratio of the vessel diameter to the impeller diameter	3.2	2.6	2.4
Tip speed (m/s) at N = 1000 rpm	2.12	3.16	-

II. Appendix: Standard Error

The standard error provides a measure of the error that exists in a normally distributed sample set. Unlike the standard deviation, the error is not calculated using the deviation from a mean value but rather differences between a pair of sample scores or residual values. The sum of the differences are squared and then divided by the number of samples. The square-root of this value calculates the standard error as described by Equation II.1.

$$SE = \sqrt{\frac{\sum_{i=1}^{n}(z_i - y_i)^2}{n}}$$

Equation II.1

Where, y and z are duplicate results analysed from a split sample and n is the number of un-split samples.

III. Appendix: Calibration of OD_{550} with Protein Concentration

A calibration curve linking the OD_{595} to protein concentration was constructed using a series of standard samples of known protein concentration. An example calibration curve is shown in Figure III.1.

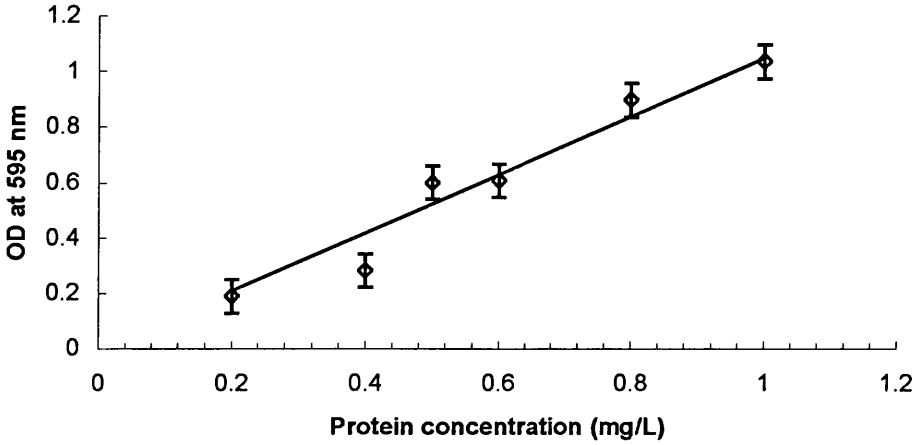


Figure III.1: Calibration of OD_{595} measurements to protein concentration for series of standard bovine albumin samples. Measurements were determined as described in section 3.1.3.2. The fitted line has a correlation coefficient of 0.95.

IV. Appendix: Standard Deviation

The standard deviation is a measure of how widely values are dispersed from the average value (the mean). STDEV assumes that its arguments are a sample of the population. The standard deviation was calculated using Equation IV.1.

$$SD = \sqrt{\frac{\sum_{i=1}^{n} (\bar{y}_i - y_i)^2}{n-1}}$$

Equation IV.1

V. Appendix: GC Calibration Curves

Concentrations of ketone and lactone in the ethyl acetate phase of samples were found using predetermined calibration curves. Figure V.1 and V.2 show typical calibration curves of solute concentration against FID response (integrated peak area) for ketone and lactone respectively.

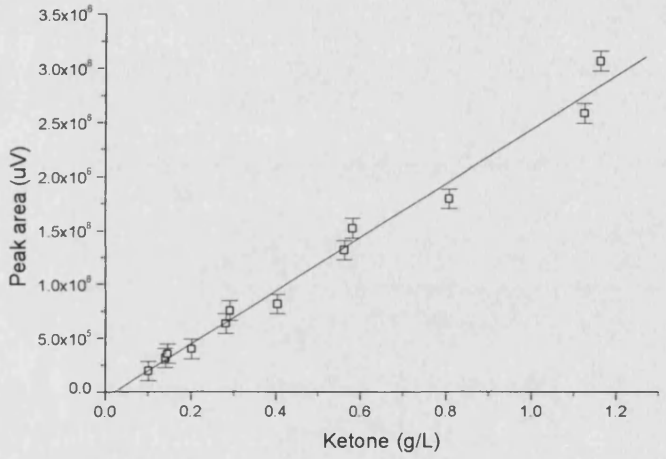


Figure V.1: Typical calibration curve for ketone in ethyl acetate. Samples were prepared and analysed as described in section 3.1.3.5. R = 0.98. Error bars show standard error of GC analysis (section 3.1.3.5.1).

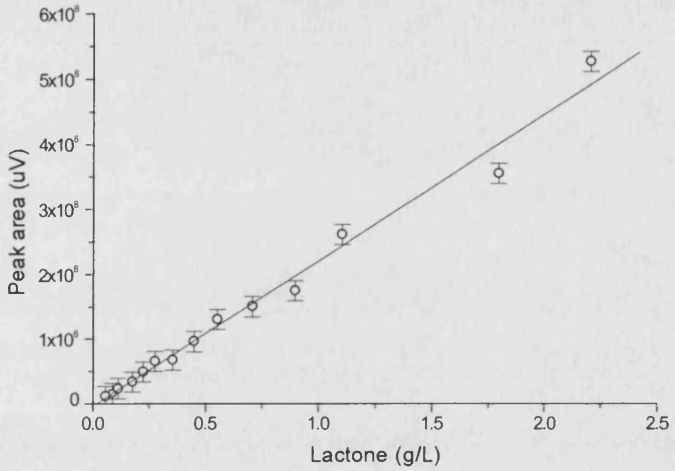


Figure V.2: Typical calibration curve for lactone[1] in ethyl acetate. Samples were prepared and analysed as described in section 3.1.3.5. R^2 = 0.98. Error bars show standard error of GC analysis (section 3.1.3.5.1).

VI. Appendix: FID response chromatogram

The flame ionisation detector (FID) response time from gas chromatography of samples prepared in ethyl-acetate is shown for ketone and combined lactone (Figure VI.1). A commercial lactone was used to calculate calibration curves for the prediction of combined lactone process samples. The commercial lactone (Figure VI.2) has the same retention time as the combined lactone (Figure VI.1).

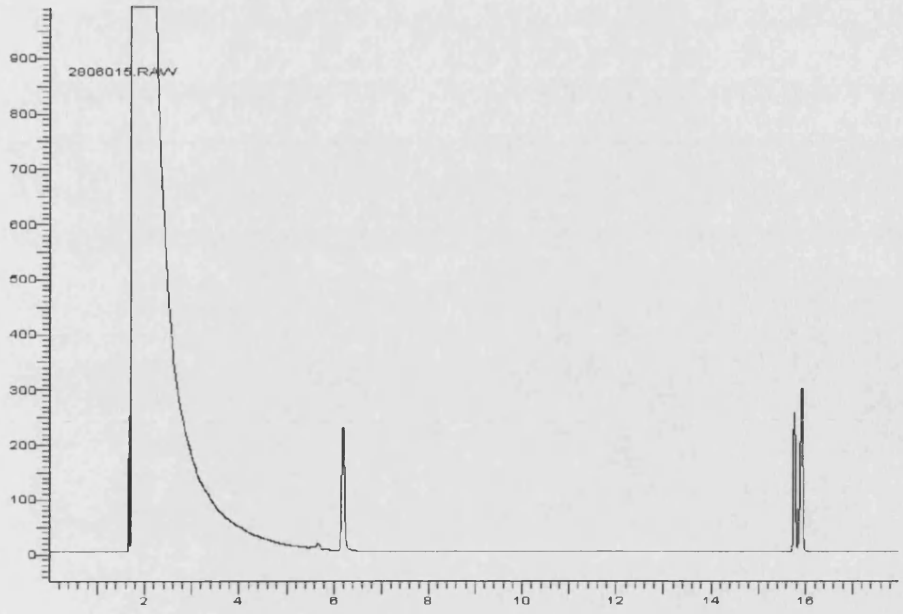


Figure VI.1. GC chromatogram of ketone and biotransformation combined lactone products.

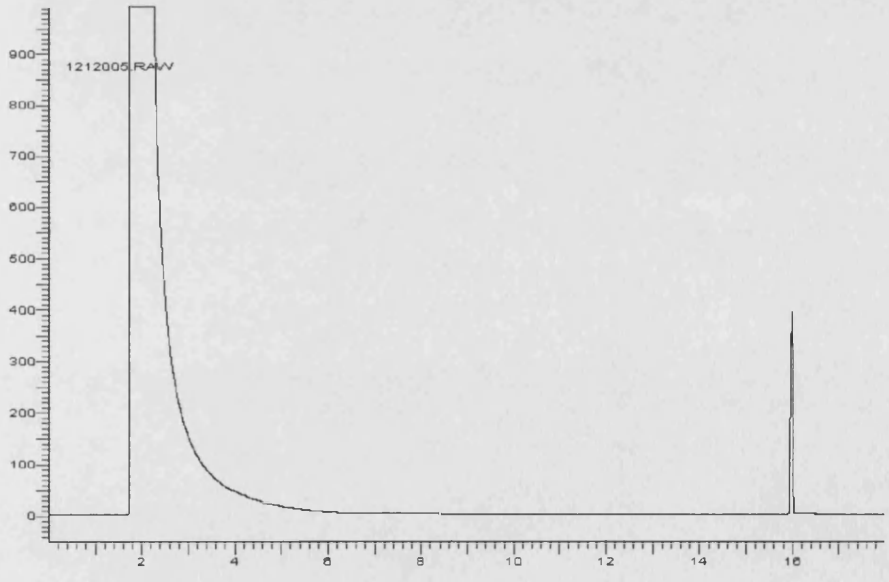


Figure VI.2. GC chromatogram of lactone[1]; 1(R), 5(S) 2-oxobicyclo(3.3.0)oct-6en-3one.

VII. Appendix: Scale-up of E.coli *Top10 [pQR239]* Batch Growth

Scale-up of agitation from the 2L to the 7L fermentation was achieved initially by means of constant tip speed (Ts). Tip speeds calculated using Equation VII.1 for LH210_a and LH2000 are given in Table VII.1.

$$Ts = \pi \cdot N \cdot di$$

N = speed of agitator (s^{-1})

di = diameter of impeller (m)

Equation VII.1

	Tip speed (ms^{-1})	
N (rpm)	**LH210_a**	**LH2000**
500	1.06	1.63
600	1.27	1.96
700	1.48	2.28
800	1.69	2.61
900	1.90	2.94
1000	2.12	3.26
1100	2.33	3.59
1200	2.54	3.91

Table VII.1: Agitator tip speeds for fermenter LH210_a and LH2000. The shaded areas depict the actual range of tip speeds used during the batch fermentation of *E.coli* Top10 [pQR239].

VIII. Fermentation Parameters

An analysis of the fermentation parameters for batch growth of *E.coli* Top10 [pQR239] is given in Table 3.2 (section 3.3.1). The Reynolds number (Re) was calculated using Equation VIII.1. Reynolds number is proportional to inertial force/ viscous force and is used in mass transfer to account for dynamic similarity. For a given geometry, the Reynolds number based on impeller diameter is convenient for describing the hydrodynamics of the system. A fermenter system with Re > 10^4 is characterised by turbulent flow where Power number (Po) is independent of Re (Rushton et al 1950). A broth density of 1100 kg.m^3 and a viscosity 0.01 Pa.s was assumed from bioprocess principles ((Doran, 1998)).

$$\mathrm{Re} = \frac{\rho \cdot N \cdot di^2}{\mu}$$

p = density (Kg.m^{-3})

$N_i d_i$ = characteristic velocity (Pa.s)

d_i = characteristic length (impeller diameter)

Equation VIII.1

The volumetric un-gassed and gassed agitator power requirement of the fermenter's was calculated using Equation VIII.2 and VIII.3 respectively. The *Pg* results are given in Table VIII.1. An impeller Po of 5 is typical for a Rushton turbine type impeller operating in the turbulent regime ((Doran, 1998)).

$$Po = \frac{Pug}{\rho \cdot N^3 di^5}$$

Pug = impeller ungassed power requirement (watts)

Equation VIII.2

$$Pg/Pug = 0.4$$

Pg = impeller ungassed power requirement (watts)

Equation VIII.3

	Pg/V (watts/L)	
N (rpm)	**LH210_a**	**LH2000**
500	0.076	0.204
600	0.136	0.356
700	0.212	0.564
800	0.32	0.84
900	0.456	1.196
1000	0.624	1.64
1100	0.828	2.184
1200	1.076	2.832

Table VIII.1: Volumetric *Pg* for fermenter LH210_a and LH2000 over a ranger of agitation speeds. The shaded areas depict the actual range agitator speeds used during the batch fermentation of *E.coli* Top10 [pQR239].

To understand better the potential effect of the respective maximum calculated power inputs for each fermenter on the *E.coli* cultures the maximum shear rate (() was calculated using Equations VIII.4, VIII.5 and VIII.6.

$$E = \frac{Pug}{\rho V} \equiv \frac{Pug}{V}$$

E = Average energy dissipation per unit mass ($W.Kg^{-1}$)

Equation VIII.4

$$\gamma = \sqrt{\frac{E}{\nu}}$$

(= shear rate (s^{-1})

< = kinematic viscosity, Equation VIII.6 ($m^2.s^{-1}$)

Equation VIII.5

$$\nu = \frac{\mu}{\rho}$$

v = kinematic viscosity, ($m^2.s^{-1}$)

Equation VIII.6

IX. Model Validation

To estimate the optimal number of factors in any given PLS calibration model a cross-validation approach was used. The PLS factor based methods find linear combinations of the independent variables (wavelength responses) which account for variation in the data set. The linear combinations are the factors, which can be considered new orthogonal variables. If in a given model an insufficient number of factors are retained in the model, future predictions are unreliable because important quantitative information is missing. However, if too many factors are included in the model, future predictions are misleading because random variation (noise) has been built in to the model. The best possible regression models were achieved by a cross validation procedure optimising the number of factors such that either a minimum prediction residual error sum of squares (PRESS) is reached (Figure IX.1) or the fewest factors required to achieve a standard error of calibration (SEC) similar to the associated gas chromatography standard error. Cross-validation involved each calibration sample being dropped out of the calibration set, generating regression equations on the reduced calibration set and then subsequently predicting the value of the dropped standard to calculate a PRESS.

To calculate the SEC on a data set the residual standard error (σ) is calculated using Equation IX.1 based on the fitted line $y = a + bx$. The residual standard error is a root mean square average of the errors about a fitted line and as such represents a typical discrepancy from the line. For the calculation of SEC, n-2 rather than n allows for the fact that the constants a and b have been calculated from the same data that are used for the errors. For the calculation of the PRESS and SEP, Equation II.1 is used. The SEC is likely to be an underestimate of the true error (SEP) since the equation used to for prediction is a line of best fit for the particular sample set. This is the reason why it is important for the sample set to represent the expected variation of the process samples.

The correlation coefficient (R^2) measures the extent to which the fitted straight-line relationship explains the variability in the y-values. It depends on the spread of the y-values as much as it does on the goodness of fit and so it measures the predictive ability of the equation relative to the range of y-values in the calibration set. It is therefore possible to affect the R^2 value by manipulating the range of values. An extended range of values might improve the R^2 whilst prediction errors actually become bigger. The R^2 value can be calculated using Equation IX.2.

$$\sigma = \sum_{i=1}^{n} (y_i - a - bx_i)^2 / (n-2)$$

Equation IX.1

$$R^2 = \frac{\sum_{i=1}^{n} (y_i - \bar{y})(x_i - \bar{x})}{\sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2 \sum_{i=1}^{n} (x_i - \bar{x})^2}}$$

Equation IX.2

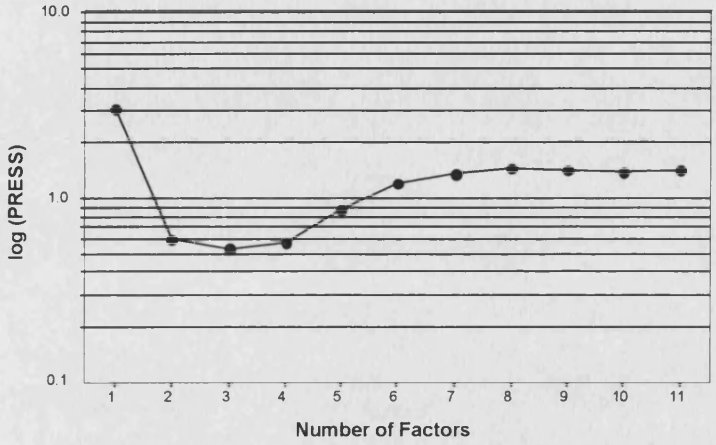


Figure IX.1: An example of PRESS analysis on ketone calibration data.

X. Rapid Bioprocess Development

London Business School

NEW TECHNOLOGY VENTURE
Group Project

RAPID

BIOPROCESS

DEVELOPMENT

Susanna Khavul
Professor - London Business School

Group Assignment - Prepared by:
Paul Bird
Marin Mignot
Sam Pickering

1. Executive Summary

1.1. Technology

Rapid Bioprocess Development (RBD) consists of accurately simulating fermentation and biocatalysis at a very small scale. Using robotics one can automate thousands of these small bioreactors at the same time, and analyse the results generated using advanced instrumentation and software. The time taken in conventional optimisation (development) and experimentation with new product candidates is largely due to the number of experiments required. Simultaneously performing these experiments allows a significant decrease in development time (equating to millions of dollars e.g. reduced time to market, extended useful patent life, higher NPV for each product).

RBD will revolutionise the biopharmaceutical and chemical bioproduct development, by bringing together three different technologies:

- Robotic liquid handling of ultra scale down reactors,
- Advanced instrumentation for monitoring, control and analysis,
- Software for modelling and directing iterative cycles of the development.

RBD is a technical breakthrough that will enable drug discovery companies to better control the risks associated with drug development.

1.2. Driving Force

Each year in pharmaceutical development means one more year before a patient can benefit from the drug and one less year of profit for its inventor before its patent expires and generic competitors jump in. As this patent clock winds down potential sales of well over $1 million are lost for every extra day spent in bringing the drug to market. The slowness in the development

process is a result of the numerous steps that a drug must pass through en route to commercialisation. Not surprisingly all players involved in the drug development process are attacking this time expenditure.

1.3. Market

There are three key markets to which RBD would be targeted:

- large pharmaceutical and chemical companies
- bioprocesses consultants and contract manufacturers
- drug discovery research groups

RBD addresses the existing market of drugs and/or bio-substances development process, and in that case, RBD will replace existing, under performing practices in the drugs development industry. Drug development costs have risen to over $600 million per drug whilst 53 leading drugs will come off patent by 2005[1]. There therefore exists pressure for the pharmaceutical and chemical industry to re-asses the route to producing new-drugs.

1.4. Commercialisation Strategy and recommendations

The technology to date is not sufficiently integrated for RBD to be commercialised without a development and incubation of the concept. However, the potential of RBD to deliver innovative solutions to current bioprocess development issues has led to the following recommendations:

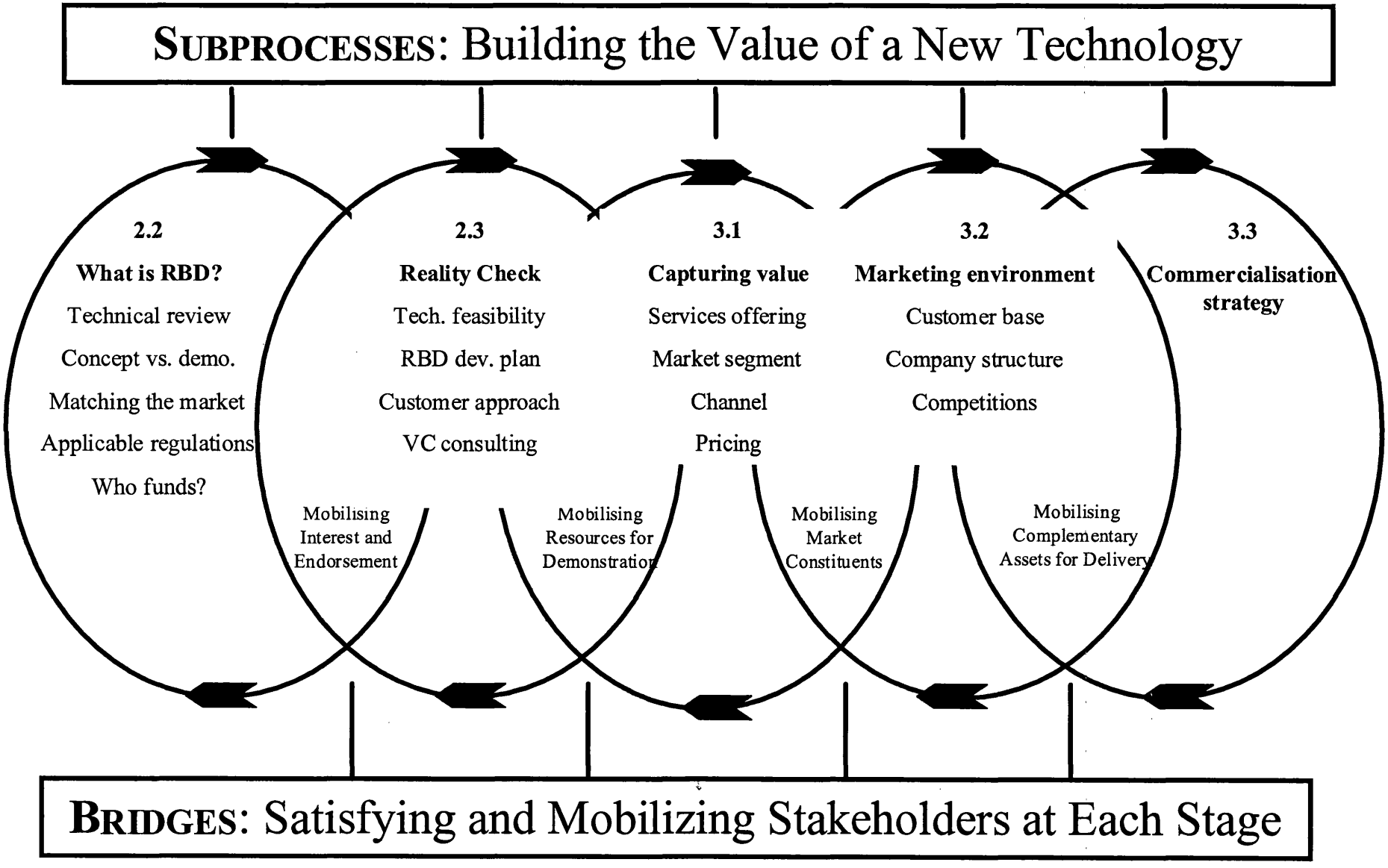
1. Integrate and develop RBD technology
    a. Expand and define capabilities of robotically operated micro-reactors
    b. Develop a near infrared analyser for 12 well and then 96 well micro-titre plates
    c. Evaluate with UCL ventures the IPR issues surrounding RBD including the business model.

d. Establish the demonstration of RBD as a milestone towards a spin-out business venture.

2. RBD product should be structured to deliver value to each customer segment
    a. Sell platform product to large pharmaceutical companies
    b. Small to medium sized biotechnology/ chemical companies a sale can include service, partnership or license agreement.
    c. RBD technology should integrate into customer development pipeline.
3. Submerge RBD into customer culture to reduce barriers to market entry
    a. Paradigm shift in drug discovery programmes should be reflected in bioprocess development phase
    b. Communicate value of RBD to customers in terms of man hours saved, expenditure saved, reduction in investment risk and reduction in time to market.
    c. Replacement of existing expertise and resources by RBD maximise existing capabilities and not challenge them.

[1] Arlington,S. Pharma 2005: the challenges.

Figure 1: THE PROCESS OF RBD COMMERCIALISATION



2. Concept definition of RBD

2.1. Overview of the Bio-industry

The simplest definition of Biotechnology is *'the manufacture of commercial products from living organisms.'* In their crudest form, biotechnological principles have been documented since 6000 B.C. when the Babylonians and Sumerians used the fermentation of yeast for the production of alcohol. It is the advent of genetic manipulation, however, that has led to massive growth in the use of biotechnology in pharmaceutical and fine chemical sectors.

Millions of years of evolution have created thousands of microorganisms containing enzymes known to catalyse almost every chemical reaction. Nature's exquisite enzymatic chemistry tends toward the highly selective, efficient, temperate and environmentally benign. The chemical industry is amassing biocatalytic expertise to meet the needs of pharmaceutical customers.

Fighting the clock: pharmaceutical drug development is slow, requiring years of research and lengthy clinical trials. The drug industry group Pharmaceutical Research & Manufacturers of America (PhRMA) claims that drug companies invest from 12 to 15 years in each new drug[2]. In biopharmaceuticals, the Tufts Center for the Study of Drug Development has determined that, while the number of new products has been increasing steadily, clinical development times have doubled since 1982 to an average of 68 months[3]. The slowness in the development process is a result of the numerous steps that a drug must pass through en route to commercialisation.

Products derived from the mapping of the human genome in 2000 are expected to start entering the market from 2003 and it is anticipated that the number of identified, tested and commercialised drugs will increase six-fold over the next 20 years[4].

---

[2] http://www.phrma.org/index.phtml?mode=web
[3] Chemical and Industry. Reinventing Pharma. January 28, 2002.
[4] http://www.biocapital.com/en/investment/trends.shtml 14:40, 17 March 2002

Each year in this protracted process means one more year before a patient can benefit from the drug and one less year of profit for its inventor before its patent expires and generic competitors jump in. As this patent clock winds down potential sales of well over $1 million are lost for every extra day spent in bringing the drug to market[5].

2.2. What is RBD?

2.2.1. Technical review of RBD

The RBD product brings together three distinctive technological developments:

1. Robotic liquid handling of ultra scaledown reactors.
2. Advanced instrumentation for monitoring, control and analysis
3. Software for modelling and directing iterative cycles of the development process

Robotic liquid handling of ultra scaledown reactors. Parallel advances in molecular biology together with the use of laboratory automation are now placing severe pressures on the process development stages of the bioproduct life cycle. The use of robotic liquid handling experimentation at microscale to obtain key process design data has the potential to overcome this bottleneck and radically improve the speed at which bioproduct processes can be established. The ability to mimic large-scale processes (>450 litres) at a microscale (<1 millilitre) has been demonstrated using a *Escherichia Coli* whole cell catalyst during a Baeyer-Villiger bioconversion[6]. In this case the microwell approach represented a 4000-fold reduction in scale and material requirements[7].

[5] Watkins, K.J, 'Fighting the Clock', Chemical and Engineering News, 27-34, Jan 8 2002
[6] Lye,G.J., Dalby,P.A., Woodley,J.M. Better Biocatalytic Processes Faster: New tools for the implementation of biocatalysis in organic synthesis. OPRD, submitted.
[7] Pickering, S.C.R, Doig, S.D., Lye, G.L., Woodley, J.M. The use of microscale processing technologies for quantification of biocatalytic Baeyer-Villiger oxidation kinetics. Biotech. Bioeng. In Press

Advanced instrumentation for monitoring, control and analysis. The use of microscale technologies for quantification of process development variables is likely to be limited by the time required to assay for products and reactants. Traditional techniques, which include gas chromatography and high-pressure liquid chromatography, are not going to be applicable if the speeds of analysis are going to lag considerably the speeds of process operation. Near infrared spectroscopy (NIR) can provide sensitive and robust monitoring of complex bioprocesses on-line (no sampling of the bioprocess is required) and rapidly (<45 seconds)[8] [9]. This compares favourably to the traditional techniques, which are off-line (sampling of the bioprocess is required) taking up to 35 minutes to complete. NIR instrumentation, however, is only commercially available in a size designed to interface with a standard stirred tank reactor port (20mm) and not a microwell (<8mm). The process of adapting the NIR instrument to interface a modified microtiter plate is necessary before incubation and then demonstration of the RBD product can be realised. Such innovation represents an opportunity to attain intellectual property rights and to add value to the RBD idea.

Software for modelling and directing iterative cycles of the development process. The software component of the RBD product can be divided into two components: i) the chemometric software, which enables NIR quantification or qualification of the bioprocess under study and ii) the software for modelling and directing the development of the bioprocess. Currently commercially available software is being used under license to perform these tasks.

2.2.2. Theoretical concept vs. initial demonstration?

The RBD concept requires a number of distinct technological developments to be interfaced. The robotic liquid handling and ultrascale down reactors (e.g. 96-well plates, see appendix 1) are

[8] Bird,P.A., Sharp,D.C.A., Woodley,J.M. Near-IR spectroscopic monitoring of analytes during a microbially catalysed Baeyer-Villiger bioconversion. OPRD, submitted.

commercially available (Perkin Elmer, Packard Bioscience) without any customisation. The knowledge and expertise required to mimic bioprocess conditions in 96-well plates using robotic liquid handling is present at the Biochemical Engineering Department of University College London. This knowledge allows a reactor of the 96-well plate (volume 400 μL) to mimic important operational variables of a pilot-scale reactor (demonstrated up to 450 L). The art of mimicking a large process at a small scale provides the basis for rapidly optimising or developing bioprocesses. There is no intellectual property rights (IPR) associated with this aspect of RBD, despite the utility and unique application of the technology. It is therefore considered that this application for robotic-liquid handling systems should form the basis of a patent search and evaluation.

The use of NIR spectroscopy to monitor and control bioprocesses has been well documented both at UCL and elsewhere. The technology is widely available from a number of vendors and for a number of applications. FOSS NIRSystems, the leading manufacturer of near-infrared (NIR) spectrophotometers (and used at UCL), designs and manufactures a complete range of rapid-scanning, near-infrared spectrophotometers for laboratory and process applications. FOSS products are designed for application in the agricultural, food, pharmaceutical and chemical industries. However, the FOSS spectrophotometer at UCL cannot be interfaced or modified for use with 96-well plates (appendix 2). Conversations with Sheelagh Halsey an Industrial NIR Applications Scientist (FOSS UK) have indicated that there is no product available from FOSS for use with 96-well plates. Sheelagh was able to indicate that development of NIR systems for use with RBD by FOSS is unlikely without market research indicating relevant commercialisation potential. Collin Dye an applications manager of Clairet Scientific Limited (UK), the specialist supplier of spectroscopic solutions to industrial problems confirmed that no NIR product is commercially available specifically for 96-well-plates. Collin was able to recommend manufacturers of products,

[9] Bird,P.A., Sharp,D.C.A., Woodley,J.M. Biocatalysis: process monitoring and control of a microbially catalysed Baeyer-Villiger lactone synthesis. BioPharm. Submitted.

which could be used to develop such a product (appendix 3). Since NIR, technology has such varied applications and since the application of NIR for RBD requires specialisation of the technology it is considered that the development of a 96-well plate NIR reader should be the focus of a development programme at UCL. This would represent an important RBD value-creating step that has clear IPR potential. Indeed a NIR 96-well plate reader could have novel utility other than that described by RBD. To facilitate this innovation three recommendations have been made:

- Current UCL NIR technology should be developed to read 12-well plates as a bridging step to 96-well plate reactors.
- Once a 12-plate NIR reader is developed, this knowledge should be applied with the necessary capital investment to the invention and demonstration of a 96-well reader.
- The 12 and 96 well NIR readers should be evaluated for there commercial potential in addition to RBD.

Pfizer Global Research the world-leading pharmaceutical company has a process analysis group. Martin Warman of the Pfizer process analysis group was able to discuss the issues for developing the NIR instrument. It was his opinion that existing NIR instruments could be quickly adapted to read a 96 or 12 well plate. Using an instrument manufacturer e.g. Spectra-Alliance, Martin believed that a NIR 96-well plate reader could be demonstrated within 3 months given the resources.

The outputs of a system that operates and monitors a bioprocess in a microwell are process specific data. This data is then applied to the optimisation and development of the process resulting in RBD. The software tools which will collect, reduce and visualise this output data represents an important aspect of the RBD technology. These software applications will allow the end-user of RBD to translate the data into valuable process specific solutions. A recommendation is therefore

made to evaluate how the data generated through RBD would best be presented to any identified customer.

Demonstration of RBD for operating and monitoring bioprocesses should become the focus of efforts once the NIR 96-well reader was available. Validation and demonstration of the RBD technology could take place to describe the utility for optimisation and development of relevant bioprocess systems. It would be desirable for a comparison of the RBD strategy to a traditional strategy to be made. The technology development plan is outlined in figure 1.

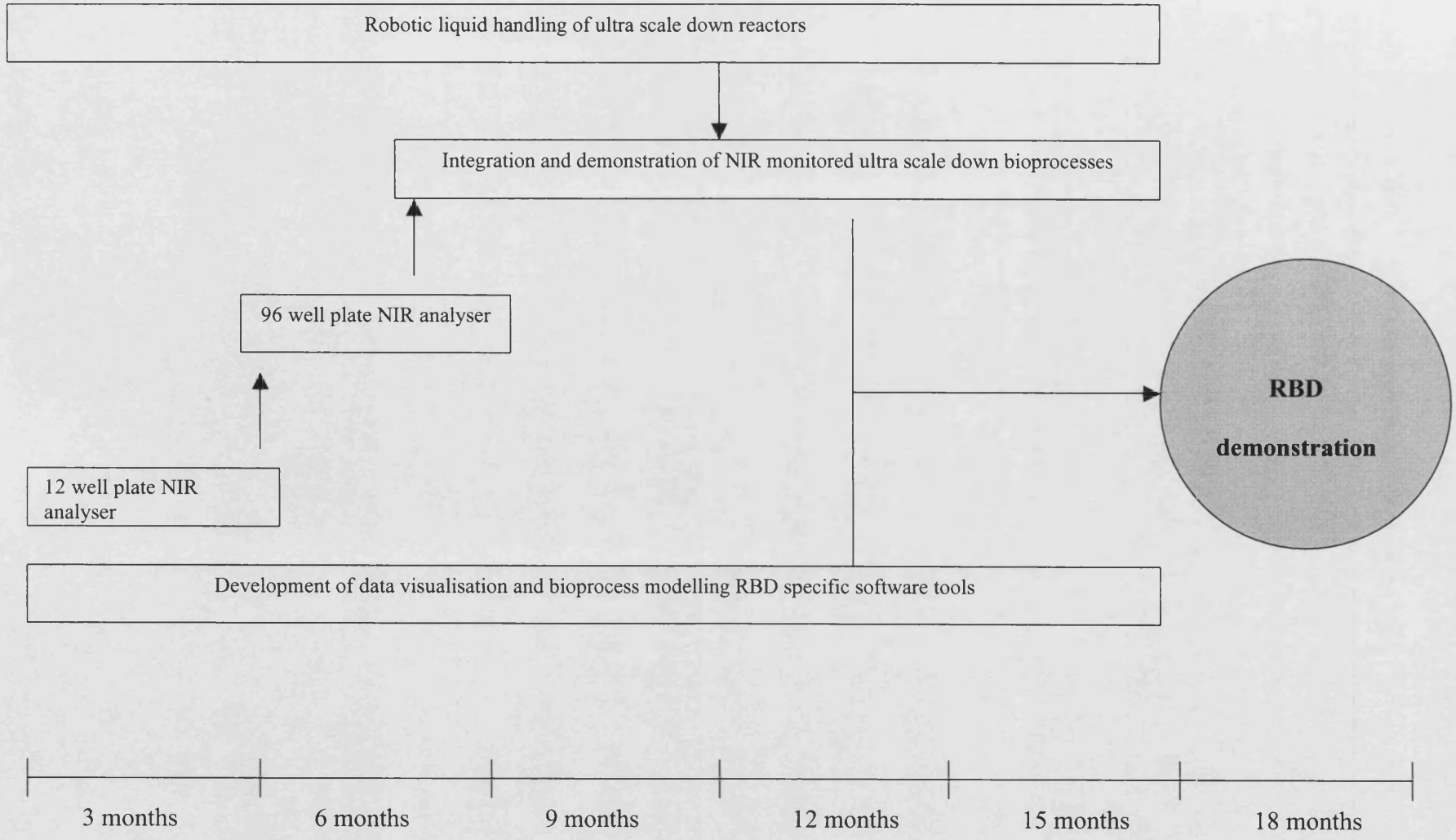


Figure 1 - Technology Development Plan

2.2.3. Intellectual Property

Although RBD integrates existing pieces of equipment manufactured by existing companies, the use of the pieces of equipment together in the manner necessary to perform RBD constitutes an innovative step and, as such, is capable of being patented[10]. The general concept of the RBD method cannot be patented, however. The disclosure of the idea in basic principles was the subject of academic grant proposals and has been discussed at academic conferences. This disclosure puts the RBD concept in the public realm and, as such is considered to be 'in the prior art' for the subject of future patenting issues.

Several patents could be written to protect both the process and the product[11]. The software written to perform the analysis, iteration and subsequent experimental design would also be subject to copyright. The algorithms developed to perform measurements from the NIR may be patentable or could be covered by the copyright on the software exploiting them. Simon Kremer recommended that the novel business models proposed by the implementation of RBD could form the basis of patent applications. The process of patenting business models has not been tested under UK law, but has been landmarked under US patent law, and is being considered here.

2.2.4. The idea: Matching the market.

There are three key markets to which RBD would be targeted:

- large pharmaceutical and chemical companies
- bioprocesses consultants and contract manufacturers
- drug discovery research groups

[10] Simon Kremer, Partner for Lifesciences Mewburn Ellis, Patent Agency
[11] Interview with Jeff Skinner, Commercial Director, UCL Ventures, Monday 11 March 2002

The average time for biopharmaceutical[12] development from the date of cloning to approval is 7.8 years, with a minimum of 4 years[13]. Where the product is a small-chemical pharmaceutical, development takes 15 years. This includes approximately six years in discovery, pre-clinical testing, and toxicity studies; one and a half years in phase one trials to assess safety in healthy volunteers; then two years in phase two trials with a few hundred patients to evaluate the drugs effectiveness and side-effects. The development process continues with three and a half years in phase three trials, then one and a half years of FDA review. Not surprisingly all players involved in the drug development process are attacking this time expenditure.

Increasing global competition, and demands for lower drug prices from governments and health care providers are putting pressure on large pharmaceutical companies to reduce these development times. RBD would allow development times to be significantly reduced, and allow earlier decision making when selecting which drug candidates to take forward.

Bioprocesses consultants and contract manufacturers could employ RBD to quickly optimise processes. This is a goal in itself for the consultant, but would also allow faster delivery of the product for contract manufacturers.

Companies or academic research groups with a single product could use RBD to demonstrate that the product is capable of being manufactured, which would add value to the product, both for proposals seeking investment, and licenses to produce the drug.

In the area of Biotechnology, Lonza Group is the world leader in exclusive custom manufacturing for the global life sciences companies. In addition to Lonza's chemical custom manufacturing, Lonza Biotec offers microbial custom fermentation and biotransformation services. Lonza Biologics is focused on the production of therapeutic monoclonal antibodies and recombinant

[12] A Biopharmaceutical is defined here as a large therapeutic drug e.g. protein molecule, antibody, vaccine.
[13] F. Foo et al., "Biopharmaceutical Process Development: Part I, information from the First Product Generation," *Pharm Technol. Eur.* 13(6), 58-64 (2001)

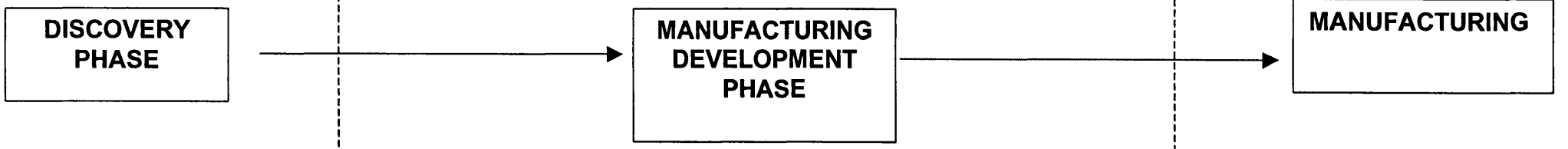
proteins produced by mammalian cell cultures. This makes them a good indicator of the consultant/contract manufacturer target market segment.

Nick Murrell[14] is currently designing a new R&D facility at Lonza Biologics. He has a keen interest in the ultra-scale down work at UCL, and is enthusiastic about the potential capabilities of RBD. In large pharmaceutical companies, manufacturing plants are generally built with a lifetime of 15 to 20 years for the production of a single entity. Lonza, however, works on short production contracts, generally lasting from a few months to a few years. This means that their manufacturing facilities are generally more flexible, and the faster they can meet the requirements of a contract, the faster they can re-engineer and re-validate their manufacturing plant. Using RBD to speed up the development of the design of the manufacturing facility conveys the obvious benefit of saving time so that there is less down-time between contracts for each facility.

A generalised bio-product development pipe-line is described by Figure 2. It can be seen that the development phase for a bioproduct is distinct from discovery but is necessary before a manufacturing phase can commence. RBD addresses the existing market of drugs and/or bio-substances development process, and in that case, RBD will replace existing, under performing practices in the drugs development industry (Figure 3). Where the discovery phase require development of a biocatalyst or bioprocess distinct from optimisation then the results of RBD will feed-back into the discovery phase (Figure 4).

[14] Nick Murrell is a Senior Process Engineer at Lonza

Bio-product Development Pipeline



Chemical and pharmaceutical companies are generating bioprocess solutions for product synthesis. The bioprocess typically includes a number of unit operations. Each unit operation requires development before the manufacturing goal can be met.

The primary objective of the development phase is to translate the bioprocess from concept to manufacturing. Development is constrained by a number of issues:

- Robust & scaleable development of unit operations
- Reaction constraints
- Fixed capital and operating expenditure
- Multiple project scheduling
- IPR
- Regulation requires any bioprocess to be fixed, after which further changes are time and resource demanding.

Production might be for R&D, clinical trials, or sale and must remain within the constraints imposed by:

- Results of each phase
- QC/QA
- Marketing
- Distribution
- Regulation
- Production ramping

Regulatory issues have the major impact on any bioprocess and approval represents a significant additional development constraint and business risk.

Figure 2 – Bio-product development pipeline.

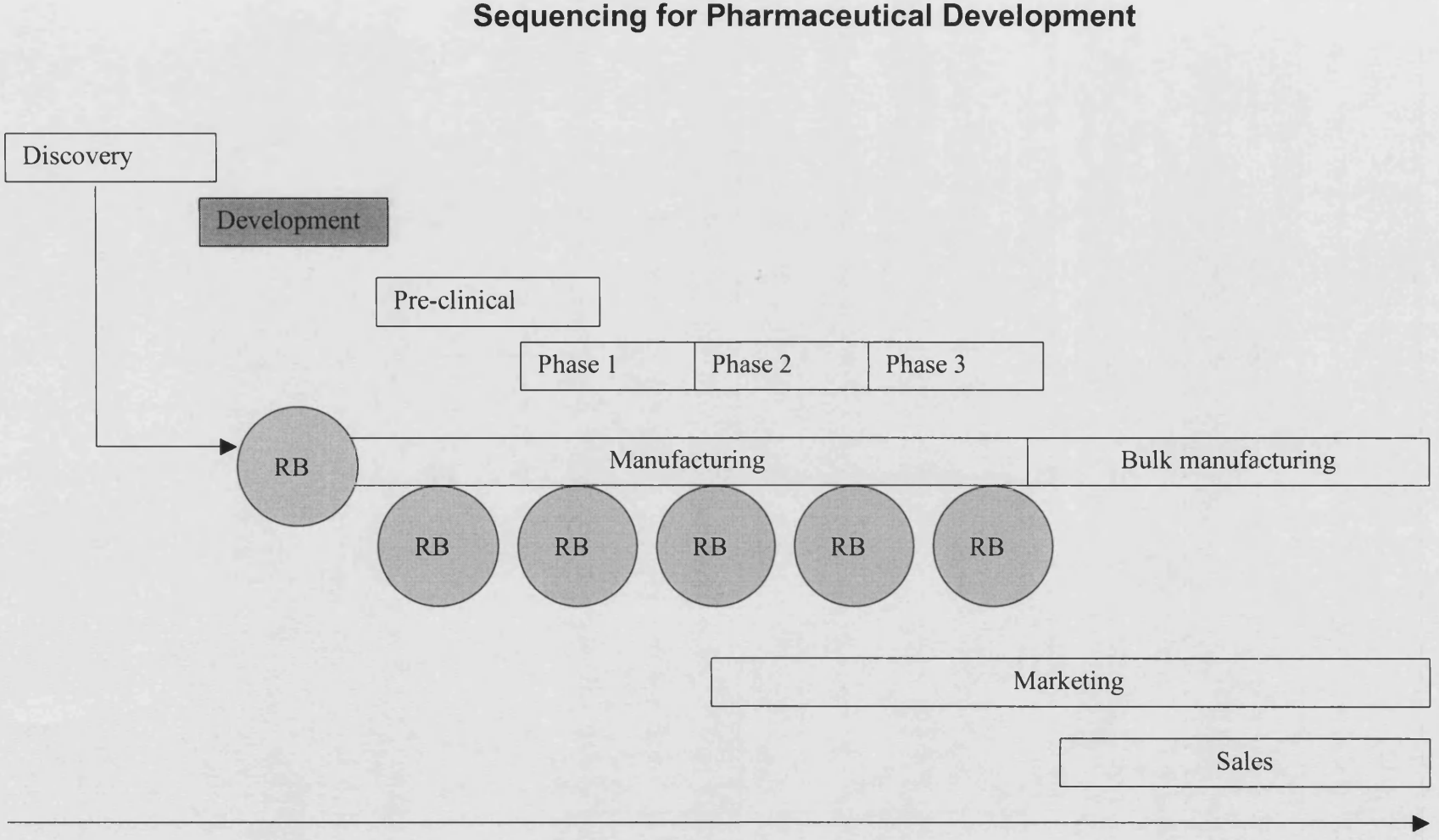


Figure 3 - Sequencing for Pharmaceutical Development

1. Sequencing for Bioprocess Development of Biocatalysis

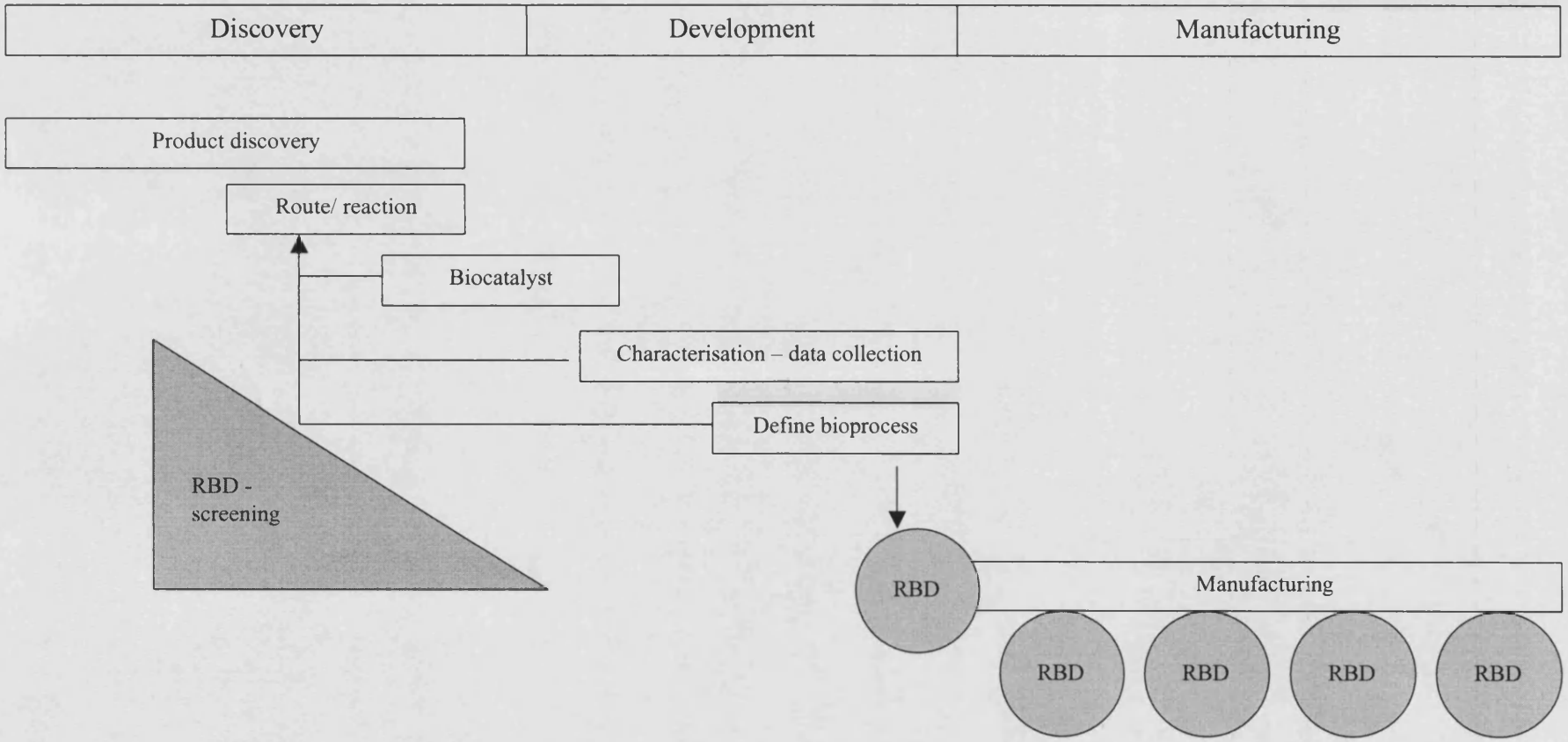


Figure 4 - Sequencing for Bioprocess Development of Biocatalysis

2.2.5. Who are the players?

The current key players are Sam Pickering and Paul Bird. They are both in the final year of researching PhDs in the department of Biochemical Engineering at UCL, working in microwell bioprocessing and automation, and NIR, respectively. UCL has a commitment to microscale bioprocessing, and is currently constructing a dedicated facility (£20M, due to open April 2002) to explore this new and exciting technology. The department has links with Packard Bioscience and Tecan (Robotic liquid handling manufacturers) and Foss (NIR equipment manufaturer), who could be consulted on integrating the necessary equipment to exploit RBD. The Biochemical Engineering Department is keen to publish new discoveries in peer-reviewed journals –this desire must be carefully controlled if the IPR of RBD is to be the subject of patent applications. The Department is a world leader in research which has made possible the production of a new generation of human therapeutic proteins. It has also pioneered research on the use of biocatalysis in the synthesis of chemical medicines. The Department manages the interdisciplinary research collaboration of ten UCL departments within the Advanced Centre for Biochemical Engineering. The Centre conducts studies on materials for human protein, gene, cell and tissue therapy and on biocatalytic processes enhanced by directed evolution and metabolic engineering. A history of patenting its research makes the department more sensitive to the needs of non-disclosure than many other departments in similar fields.

2.2.6. Applicable regulations.

The cost of bringing complex drugs to market is increasing, impacting on both the drug development pipeline and return on investment for drug development companies. The regulatory measures imposed by governments are hampering the long-term profitability of the pharmaceutical industry. Regulatory authorities are insisting that good manufacturing practices (cGMP) are

applied at a much earlier stage than previously required; the time available for bioprocess development has therefore been reduced further[15]. The development phase of drug development requires no specialist regulation other than those that are required for containment and health and safety, however the requirement for early cGMP procedures and the consequential time constraints distinguish further the advantages of RBD. Likewise, where the bioprocess under development is not a drug any regulation applicable to the manufacturing process will be associated with the end-use and are therefore unlikely to affect RBD operation.

2.2.7. Who funds?

The current research by Sam Pickering and Paul Bird is financed by the Biochemical Engineering Department, the Biotechnology and Biological Sciences Research Council (BBSRC), and Pfizer. The funding covers salaries, premises, equipment costs and lab consumables. The rights to income from any patents stemming from this research would be divided, subject to negotiation, amongst the two researchers, UCL, and Pfizer.

2.3. From reality check to market.

2.3.1. Technical feasibility vs. commercial potential?

As has been mentioned it is considered that an NIR instrument capable of monitoring a bioprocesses in 96-well plates could be demonstrated within three months. If this development was conducted within UCL for IP and resource issues then this time might increase. *Comment form David Walker UCL Optics group.* It therefore remains for a demonstration of the RBD product to be carried out in the context of the customer. This would include a detailed analysis of what bioprocesses can be developed using RBD. Are there particular bioprocesses, which cannot be

[15] Foo F. *et al.* Biopharmaceutical process development: Part 1, Information from the first product generation.

operated using micro-well reactors? This evaluation should also include expected bioprocess conditions e.g. pH, temperature. A potential customer might be a manufacturer of protein bioproducts. Bio-products typically require a fermentation step. This fermentation step involves the growth of single-cell organisms both eukaryotic and prokaryotic. The question remains as to which organisms can a 96-well plate bioprocess operation be applied.

Current research in the department of Biochemical Engineering is resolving the issues associated with the control of pH, temperature and sterility in the microwell environment. Ten hour fermentations have been carried out to exactly mimic those at a larger scale, but longer fermentations with a need to control pH and feeding are proving to be more problematic. It is anticipated that these problems will have been resolved by the end of 2002.

Even in its current state the usefulness of RBD has been demonstrated in biocatalysis with no fermentation step. Microscale processing techniques have been shown to be a useful tool for the rapid and efficient collection of biotransformation kinetic data and as a basis for bioprocess design. Automated liquid handling systems can reduce labour intensity while the small scale reduces the demand for scarce materials such as reactant, product and biocatalyst. We have demonstrated that the microscale processes in RBD can be used for the collection of quantitative kinetic data that is essential to robust development of a bioprocess[16].

2.3.2. RBD program development.

This development plan for RBD recognises that a significant investment of time and resources is necessary before a RBD product can be demonstrated but that at each stage of development value is being created and that RBD represent only one way in which this value can be realised. Figure 1

BioPharm, June 2001.

[16] Pickering, S.C.R, Doig, S.D., Lye, G.L., Woodley, J.M. The use of microscale processing technologies for quantification of biocatalytic Baeyer-Villiger oxidation kinetics. Biotech. Bioeng. In Press

depicts a recommended strategy for building the value of RBD and figure 2 depicts the process of technological development.

This program reflects the lack of IPR associated with the individual components of the RBD technology and the overall concept. A conversation with Simon Kremer a patent attorney with Mewburn Ellis highlighted the lack of consideration that has been given to IPR with regard to RPD. It was Simon's opinion that in principle a patent could be filed based on the novel utility of the technology for RBD. It was impressed that the risk of disclosure could complicate a patent application and an early patent file would be advised. That is the technological detail of the RBD technique should be clearly stated and the process of acquiring IPR should be instigated through UCL ventures. Since it is only necessary to demonstrate technological feasibility and no demonstration is required to file a patent, it is recommended that the inventors discuss with UCL and UCL ventures the IPR issues surrounding RBD.

2.3.3. Customers approach.

The three target markets for RBD, as defined in section 2.2.4, encompass huge number of companies. In order for RBD to be accepted, the large companies (Figure 5) listed in the table below, should be approached first.

Large Pharmaceutical	Contract Manufacturer	Research with no manufacture
AstraZeneca	Lonza	Academic research groups
Bristol-Myers Squibb	Bayer	Drug discovery
GlaxoSmithKline	DSM Fine Chemicals	Biotechnology companies
Aventis	Solutia	Research institutions
Merck & Co., Inc.	Charybdis Technologies, Inc.	

Pfizer Inc.	Chiron Contract Manufacturing	
Roche Bioscience	Q-One Biotech	
Eli Lilly	Rhodia Chirex	
Novartis	Molecular Medicine BioServices, Inc.	
BMS	Formatech, Inc.	

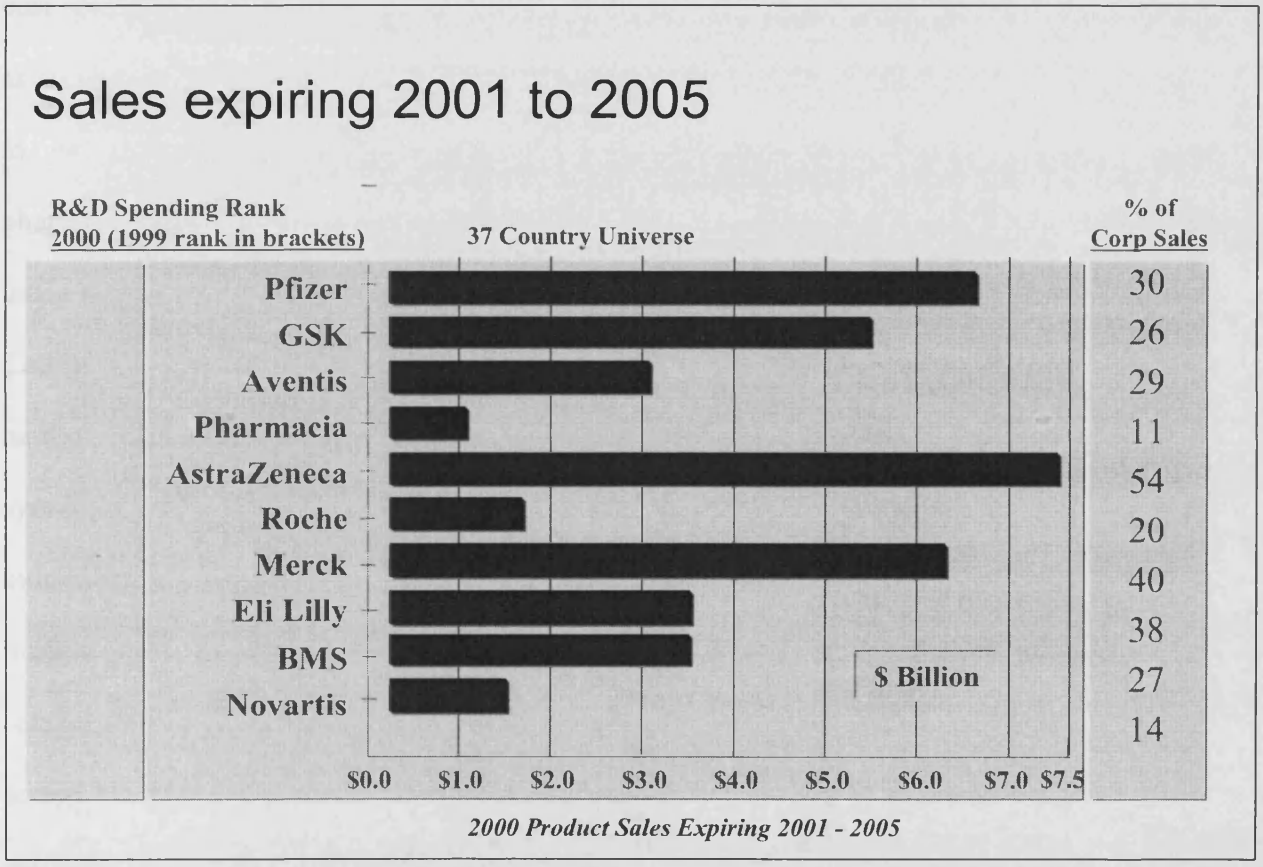


Figure 5 – Steve Arlington, Pricewaterhouse Coopers, Pharma 2005 Presentation

The value of RBD to the customer will reflect the application of RBD during product development. It should be noted that there are two distinct utilities of RBD (please refer to section 2.2.4). If RBD is applied to manufacturing development (figure 3) the customer will potentially achieve using RBD an optimum solution for manufacturing, where previously a compromise

solution was only possible due to resource and time constraints. In addition, the optimum solution can be gained in a significantly shorter time. The value of RBD to the customer in this instance will include:

1. A reduction in the time to market for the product of the bioprocess
2. A systematic and robust approach to bioprocess development

Where RBD is applied to biocatalyst development (including screening – figure 4) then the customer could use RBD to characterise the biocatalyst. Biocatalysts are increasingly been sought as an important route to compound synthesis, especially where the product of a reaction is required to be optically pure. Biocatalyst research discovery groups in both the chemical and pharmaceutical industries screen libraries of biocatalysts for specific characteristics including those detailed in figure 2. One such biocatalyst library has recently been launched by Maxygen a California based biotechnology company[17]. Merck's approach[18] to introducing a biocatalytic step in the synthesis of small pharmaceutical molecules is initiated by the discovery of a specific synthetic step where a biocatalytic reaction could be of interest. Once it is established that a biocatalyst would present technical or financial benefits, decisions of technical feasibility are then made based on internal expertise. Procedures for screening microorganisms or enzymes that catalyze the desired step are then designed and implemented. After successful completion of screening, a process is developed at laboratory scale and later scaled-up in the pilot plant. The value of RBD to a customer who wants to characterise and develop a library of biocatalysts would be attributable to

1. The efficient and cost-effective synthesis of the product.

[17] http://www.codexis.com/frames.htm
[18] Chartrain et al., "Asymmetric bioreductions: application to the synthesis of pharmaceuticals", Journal of Molecular Catalysis B: Enzymatic 11, 502-512 (2001)

2. Process development solutions that are conceived within the discovery phase. (i.e. manufacturing issues are included in the discovery phase – see Figure 2).

True quantitative evaluation of the value of RBD to the pharmaceutical and chemical industry is difficult to gauge, as RBD is a concept from which practical utility is being foreseen. Tony Harrop, a leader of a bioprocess development group at Pfizer Global Research signalled strong approval of the RBD concept when he asked, "when can we have it?". Although Tony Harrop could not value RBD quantitatively, he was able to confirm that there would be a significant competitive advantage to any pharmaceutical company, which could shorten and or reduce the risks of product development. He detailed a scenario where the launch of a product could have been delayed due to difficulties with manufacturing and that RBD could have provided an opportunity to reduce this risk.

Drug development costs have risen to over $600 million per drug whilst 53 leading drugs will come off patent by 2005[19]. There therefore exists pressure for the pharmaceutical and chemical industry to re-asses the route to producing new-drugs. Steve Arlington of Pricewaterhouse Coopers has published four targets, which the pharmaceutical industry needs to achieve to sustain total shareholder returns by 2005:

- Double number of drug leads with improved quality of candidates entering development.
- Double success rate in development with near zero attrition after phase II.
- 3 years additional peak sales with accelerated late development and launch.
- 20% increase in revenues per drug.

As RBD aims to save time in the development phase and to improve the quality of bioprocess solutions, there is an obvious contribution to be made by RBD towards meeting these targets.

2.3.4. VC consulting.

The equipment involved in carrying out RBD is relatively costly: a liquid handling robot costs £150,000 and an NIR system costs £43,500. UCL currently owns two liquid handling robots, an NIR system, and other costly microwell measurement devices, and has funding to make other significant purchases as part of a Joint Infrastructure Fund (JIF) award. At this early stage it makes sense for the RBD project to exploit these resources, rather than seek investment to purchase dedicated equipment. The UCL owned equipment could also be used for early commercial contracts, before additional investment is requires. Jeff Skinner from UCL Ventures[20] suggested that potential Venture Capital investors would be reluctant to fund a company that used UCL owned equipment as it is likely that it would only be available for commercial use when it was not being used for academic research. A likely financial model, therefore, would be to use the UCL owned equipment for the development of RBD, testing, and early contracts. On the basis of success up to that point, investment could be sought to supplement income form early contracts and buy dedicated equipment and premises to spin-out an RBD company.

3. RBD entry to market.

3.1. Capturing value.

3.1.1. RBD Services

We have developed a method of accurately simulating fermentation and biocatalysis at a very small scale. Using robotics we can automate thousands of these small bioreactors at the same time, and analyse the results generated using advanced instrumentation and software. The time taken in conventional optimisation (development) and experimentation with new product candidates is largely due to the number of experiments required. Simultaneously performing these experiments

[19] Arlington,S. Pharma 2005: the challenges.

allows a significant decrease in development time (equating to millions of dollars e.g. reduced time to market, extended useful patent life, higher NPV for each product). Figure 6 shows the key stages in development for a large pharmaceutical company. RBD could be used at each of the development phases, speeding up conventional development times.

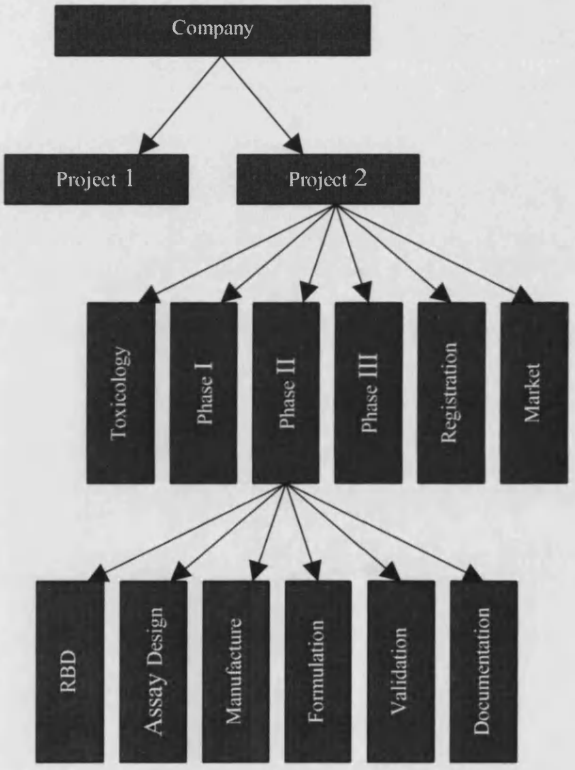


Figure 6 – Key stages in development process for a pharmaceutical company

3.1.2. Market environment.

Pharmaceutical companies are notoriously sluggish in adopting new technologies – this may be thought to be a barrier to entry into the market. Demonstrating that RBD works would only be one hurdle – if the R&D and Manufacturing departments can produce the product in a profitable form,

[20] UCL Ventures is the technology transfer office responsible for the management, protection and commercialisation

without having to rely on new/unproven technology, then that is seen as preferable. An example of this is large scale mammalian cell culture which is only just starting to be adopted by Merck, despite having been present in academic laboratories for over a decade. Previous requirements were met by duplicating the small-scale roller-bottle equipment in large warehouses, as demand was low. It is only the necessity to produce large amounts of mammalian cells for their new HIV vaccine that is forcing them to adopt large scale culture in their future plans[21].

3.1.3. Marketing Strategy

In developing RBD marketing strategy, we have first segmented the market and defined the value proposition RBD will offer to each segment, now we are going to concentrate on the promotion strategy. It is all about communications and advertising management, the objective is to ensure that all the elements of the communications mix are coordinated with the marketing strategy, delivering the same underlying positioning and value proposition to the target segment.

RBD shall be promoted emphasising the need for changes in the drug development process, the pressure is strong to save costs and reduce time-to-market, which is exactly where RBD brings value. The promotion of RBD shall be focus on making RBD a fundamental element in the redesign of R&D models, these new models being necessary to sustain the pharmaceutical and chemical companies' shareholders value.

R&D productivity is progressively declining; the amount of money spent per drugs delivered to the market is increasing dramatically:

- ✓ 1990: $20bn spent in R&D for 45 NCE[22]'s launched per year,

of UCL's intellectual property.

[21] Barry Buckland, Vice-president R&D Merck , 14 March 2002

[22] NCE = New Chemical Entity

- ✓ 2000: $45bn spent in R&D for 25 NCE's launched per year,

This trend is participating in the decrease in shareholder returns, we can observe:

- ✓ R&D costs continue to increase at 10.8% p.a., where the average growth market is about 8.1% p.a., contributing to the erosion of margin and though diminishing value,
- ✓ Revenue per new drug launched averages $265 p.a., also in constant decrease,
- ✓ Revenue from the existing portfolio diminishes at 5% p.a. due to patent expiries and price constraints,
- ✓ Total revenue from drugs grows at 7% p.a.

RBD will enable the development firms to reverse this trend, by bringing more drugs to the market, quicker and cheaper. For a comparable R&D budget, RBD enables a significant gain in productivity and consequently better off the bottom line.

The other argument to be put forward is the limited number of blockbusters, fewer than 4% of products generate sales of $500 million or more, see table:

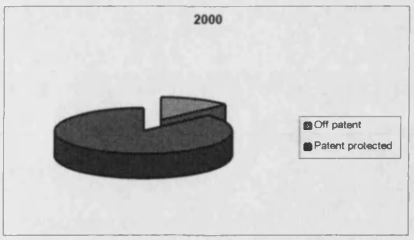
Sales Total per annum ($)	% of NCE's
1.8bn – 3.6bn	1%
920m – 1.8bn	1%
460m – 920m	2%
180m – 460m	6%
<180m	90%

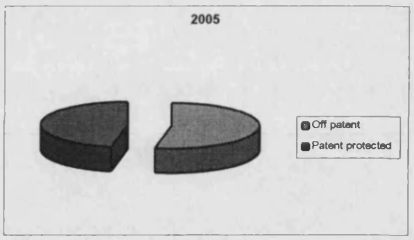
***Source**: Scrip Papers, PJB Publications Ltd 1995, Adjusted for inflation to 1998*

As RBD will enable a systematic and robust approach to bioprocess development, the odds are it will enable to higher the percentage of money-making drugs.

Patent exposure is another source of promotion for RBD:

- ✓ In 2000, 13 of the top 100 drugs, with 1998 worldwide sales of $11bn came off patent and generic competition intensified,
- ✓ In 2005, 53 of today's top 100 prescription drugs will be off patent, with 1998 worldwide sales totalling $53bn.





Additionally to the budgetary constraints, to which RBD can play a great role, these are some social constraints for new development process, which can be a strong marketing vector for RBD:

- ✓ Pharmaceutical drugs remain the 6th largest cause of death in the US,
- ✓ On average, only 40% of individual benefit from a particular drug,
- ✓ Drug development costs have risen to over $600million par drug, meaning the costs of PIII mega-trial failures is significant,
- ✓ Many medical needs remain unmet,
- ✓ Regulators struggling with risk versus benefit equation.

Again all these factors contribute at making RBD a necessary ingredient of the new development process, for a new model to produce safer, more efficacious and cheaper drugs. In defining the marketing and promotion campaign of RBD, these issue will be put forward the raise market awareness of RBD and the advantages it can bring to its customers.

It is important that any promotion reflects the culture of the customers being targeted. Large pharmaceutical companies make development investment decisions based on the business risk of investing in a given project. Appendix 4, a Gantt chart describes the risk investments made by Pfizer Ltd during the development of an unspecified product. This information, provided by an important potential customer allows an insight into the process of decision making. Each risk investment phase (appendix 4) is broken down into project stages over time with the risk (to the business) described by the expenditure and number of man hours used. As the number of risk investment phases increases so does total expenditure and the number of man-hours invested. The example given in appendix 4 required a total expenditure of $16.5 million and 119 man (person) years. The approach taken to promote RBD should communicate the value to the customer in

terms of person-years saved, and reduced expenditure. This translates to an overall reduction in investment risk at each stage of pharmaceutical or chemical development.

3.1.4. RBD delivery to the market.

The way in which RBD is delivered to the customer will reflect directly the method in which the value created through incubation and demonstration of the RBD technology is best realised. Given the market segmentation of large/ small molecules, large/ small customers and manufacturing optimisation/ biocatalyst development (or screening) there is potential for a diversified customer base. The method in which each market segment is served by the RBD business requires further consideration. Large pharmaceutical companies are unlikely to consider a licensing agreement through which RBD provides a manufacturing development service. Tony Harrop of Pfizer Global Research agreed with this analysis. Pharmaceutical companies have a culture of expertise, which might be undermined by a licence agreement that includes provision of a service. A UCL start-up, Ensynthase, discovered that problems of undermining expertise, sharing information and fixed internal budgets were all issues when trying to tender business with pharmaceutical companies. The delivery of RBD to the pharmaceutical industry would therefore be best served by sale of the technology and operational expertise.

Small to medium sized customers including biotechnology companies might use RBD to add value to their product candidates (section 2.1.a) before seeking a development partner. The customer is unlikely to want to purchase the RBD technology/ expertise due to a lack of resources, time and practical requirement. In this case, a licensing or service agreement should be considered where the business of RBD includes delivering a practical development solution or consultancy. An option would be to collaborate with companies, to create better bioprocesses and products of

bioprocesses. If RBD retains a right to royalty from the developed bioprocess and or product then a potential for high returns similar to that of any biotechnology product exists.

Servicing both segments would require a significantly different approach that would be reflected in any business plan for RBD. At this stage, it is recommended that no decision be made until IPR and a more precise evaluation of market size and value can be conducted.

3.1.5. Pricing.

The price will match the perceived value customers are ready to pay for RBD. Given that there is no integrated direct alternative to RBD then the price should reflect the time saving or value obtained through RBD. Where a RBD product is sold to a customer the price should reflect not only the cost of the technology but the knowledge base required to mimic large bioprocesses at a microscale.

3.2. Marketing environment.

3.2.1. Who are the customers?

All major potential customers, as identified in section 2.3.3, and selling approach (section 3.1.3) are defined in order to conquer the market as the product nears this stage of commercialisation. When this stage is reached we should structure our customer base, by planning the customers approach.

3.2.2. Company structure.

Presently there is no company and it has been recommended that RBD development be continued until demonstration within the research environment. After demonstration, seed funding from

Bloomsbury Bioseed or SPIM[23] could be used to spin-out a company. This would be an important milestone towards building sufficient value and potential that a significant investment round could be sought. Discussion with Jeff Skinner at UCL ventures suggested that the structure of the company reflect the ambitions for further investment.

3.2.3. What about the competition?

Current competition in the field of ultrascaledown is limited to two University groups:

1. Biochemical Engineering, Saarland University, Saarbrucken, Germany[24]

People: Gernot. T. John, Christoph. Wittmann, Prof. Elmar Heinzle

Projects:

1. Microplate cultivation of Organisms with Optical Measurement of pH and dissolved oxygen

2. Mass Spectrophotometric Respiratory Measurements in Microwell Fermentations

3. Mass Spectrophotometric Respiratory Measurements in Mammalian Cell Culture Microreactors

4. Microreactor Screening for Biochemical Conversions

5. Screening of Biochemical Kinetics Using Mass Spectrometry

6. Metabolic Engineering

2. Department of Biochemical Engineering, Aachen University of Technology, Germany[25]

People: Tibor Anderlei, Robert Hermann, Stefan Lotter, Ulrike Maier, Jochen Buchs

Projects:

[23] Sussex Place Investment Management
[24] http://www.uni-saarland.de/fak8/heinzle/
[25] http://www.biovt.rwth-aachen.de

1. Characterization of microtiter plates as small scale culture systems (oxygen transfer, evaporation...) for biological screening

2. Detection of pH in small scale culture systems with immobilized fluorescence dyes

3. Quantitative optical measurement of hydrodynamics in small scale systems

4. power consumption, small scale bioreactors, flowsheet optimization

5. Fermentation under increased head space pressure

6. Development of a novel milliscale bioreactor system for continuous (multistage) operation

7. Characterization of the specific power input and gas/liquid mass transfer in a self priming aeration system

8. Experimental investigation, modelling and optimisation of the maximum oxygen transfer capacity in shaking flasks and shaking test-tubes

9. Design, analysis and optimisation of shaken screening-systems

10. Influence of foaming on the gas-liquid mass transfer in baffled shaking flasks

11. Investigation of the hydrodynamics in shaking flasks

12. characterization and optimization of fermentation conditions in small scale cultures (test tubes, shaking flasks)

13. development of fermentation strategies for the production of recombinant proteins with different microbial systems

Neither of these research groups is thought to pose significant threat, due to the prohibitive costs of financing equipment purchases on conventional academic grants.

Currently bioprocess development is undertaken in house at large scale or using contract manufacturers and consultants. These companies have established links with the industry and are highly networked, but are themselves potential customers of RBD. RBD represents a paradigm shift in the approach that would be taken by these competitors, who are currently not delivering the same level of service that is proposed by RBD.

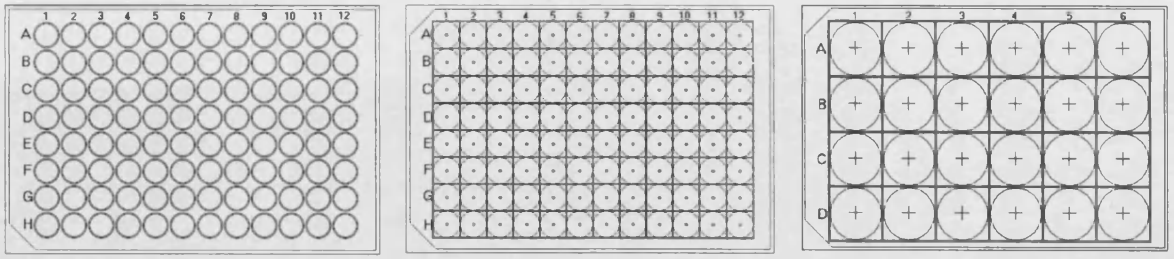
3.3. Commercialisation strategy next step for RBD.

To conclude, we believe we have demonstrated through the report the commercial potential of RBD, the market it addresses, the value it brings to its customers and the way to deliver its value proposition. The next step is to lay down a business plan to leverage RBD potential to a full market application, once the latest technical development have been completed; this could be achieved following our recommendations.

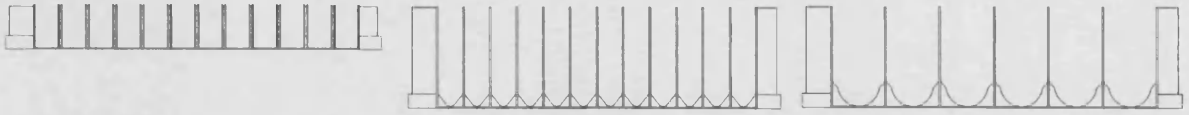
RBD is on its way to revolutionise the drug development industry.

Appendix 1 a) – Microwell Plate Technologies

Plan

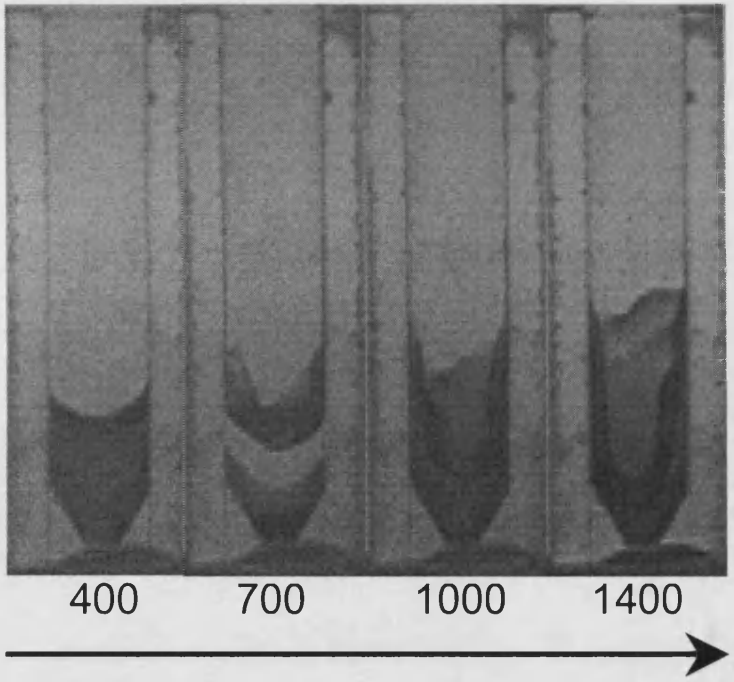


Cross-section



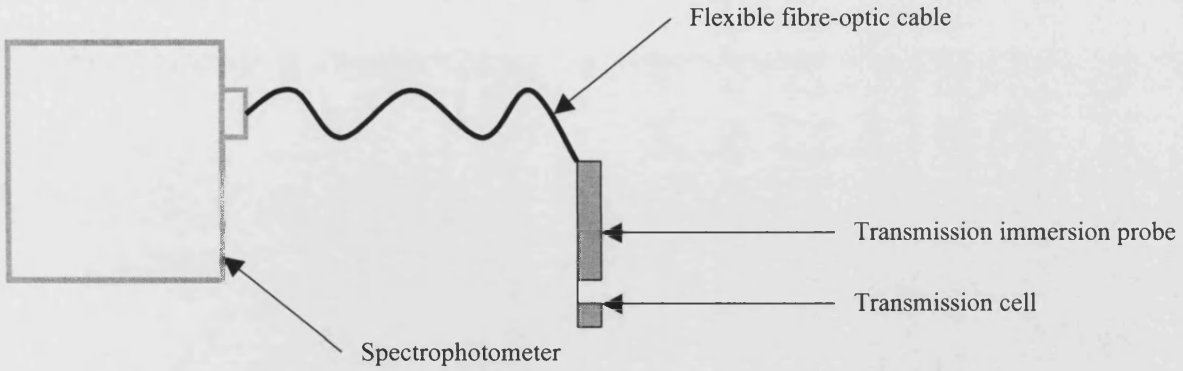
Well Type	96 Standard Round	96 Deep Square	24 Deep Square
Width (mm)	6.5	8.0	17.0
Height (mm)	11.0	37.0	37.0
Volume (µL)	360	2300	10000
Surface Area (mm^2)	33.2	64.0	289.0
Bottom Shape	Flat	Conical	Hemispherical

Appendix 1 b) – Microwell shaking demonstration (high speed video pictures)



Increasing RPM

Appendix 2 - FOSS NIRsystems spectrophotometer model 6500.



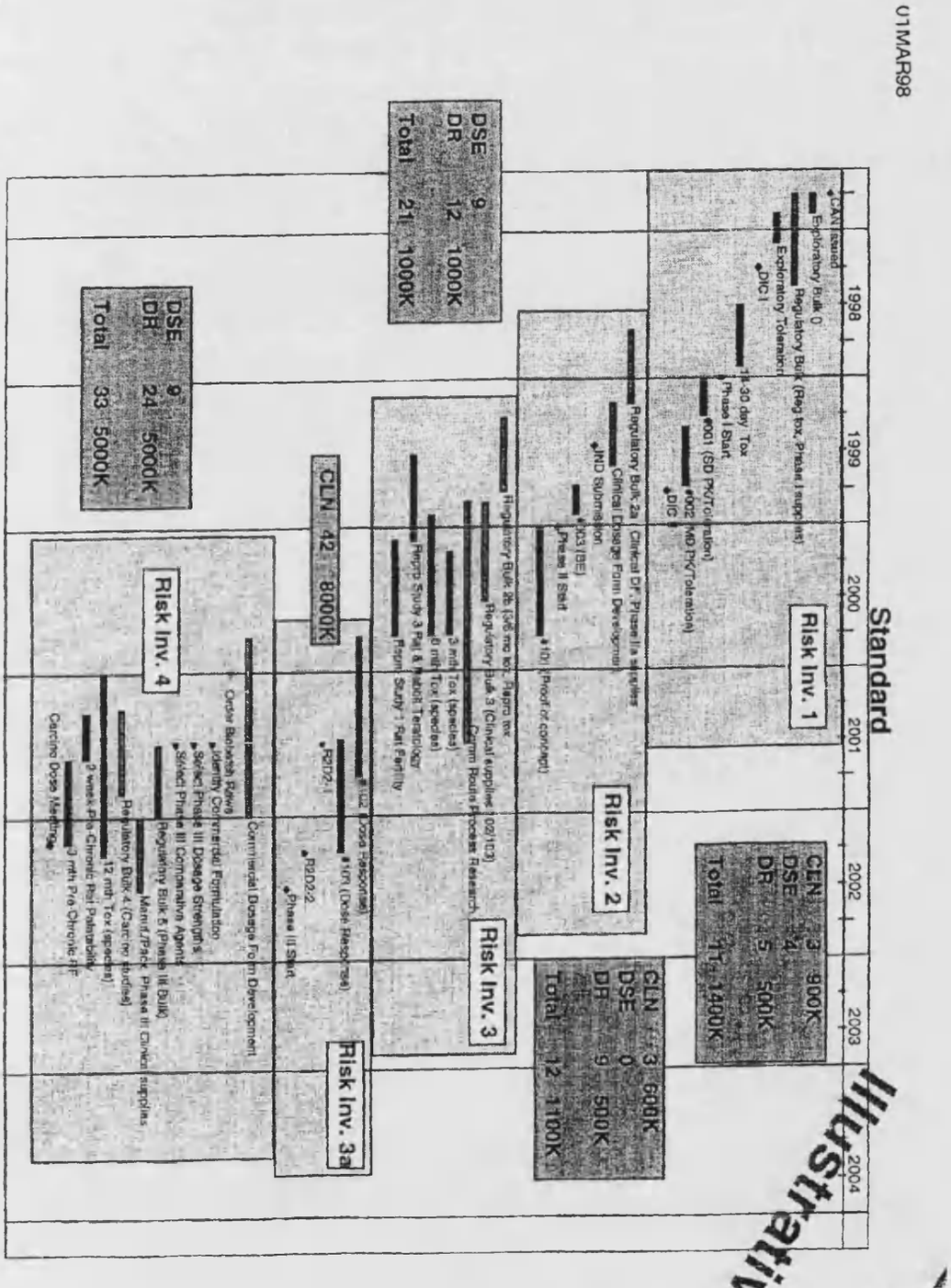
The FOSS NIR analyser transmits near infrared radiation along fibre optic cables to a transmission probe. The probe is immersed in a liquid sample of the bioprocess so that the transmission cell is covered. The sample of a bioprocess absorbs the radiation such that the transmitted radiation detected by the spectrophotometer represents that sample at that time. The FOSS fibre-optic probe attached to the model 6500 spectrophotometer has an outside diameter of 21 mm and displaces approximately 5 ml of sample. It is therefore not suitable for sampling of microwells of a 96 well plate, which have an internal diameter of 6 – 7 mm and typically hold a volume of 0.5 ml.

Appendix 3 – *96-well plate NIR transmission analyser*

There are a number of technological developments required before a 96-well plate NIR analyser can be demonstrated. The primary requirement is the procurement of a suitable NIR analyser with the potential to interface a transmission probe to a 96-well plate. Collin Dye of Clairet Scientific was able to recommend two instruments (Zeiss MCS 500, Germany and ABB Bomem, Canada) capable of being fitted to mono-fibre transmission probes. FOSS NIRsystems also produce a single fibre analyser. A quote from Varian analytical systems (UK), for the Varian Cary 500 spectrophotometer (NIR analyser) came to £40,800. This is a significant capital investment, which if funded through a research grant would take typically 12 months to arrange. A research grant would be funded through a research council and there is no guarantee of success. Dr.P.Dalby of UCL, Biochemical Engineering Department puts the chances of a research proposal being accepted at 40%.

To the NIR analyser a suitably modified transmission probe of <6mm would be fitted. A study of products on the market has indicated that Hellma Worldwide, Germany could supply this technology. A quote from Hellma for such a NIR probe states that the cost of such a device would be £2500. The conversion of the FOSS model 6500 NIR transmission probe to read a 12 well plate would not require purchase of a new spectrophotometer or probe. It is therefore recommended that a 12 well plate NIR analyser be demonstrated before development of the 96 well analyser.

Appendix 4 – risk investment phases for development of an unspecified product at Pfizer Global Research



CLN = Clinical Trials or Toxicology StudiesDR = Developmental Research
DSE = Drug Safety and Efficacy

e.g. CLN 3 900K relates to clinical, 3 man years, $900, 000 expenditure